Part 2 of the Short Course

2000 Nuclear and Space Radiation Effect Conference

# Optoelectronic Devices with Complex Failure Modes

Allan Johnston
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA), under the NASA Electronic Parts and Packaging Program (NEPP), sponsored by Code AE.

# Part 2
# Optoelectronic Devices with Complex Failure Modes

## *1 - Introduction*

This part of the NSREC-2000 Short Course discusses radiation effects in basic photonic devices along with effects in more complex optoelectronic devices where the overall radiation response depends on several factors, with the possibility of multiple failure modes. Complex failure modes can occur either because of the way different types of components within an optoelectronic structure are related (including physical factors that affect light transmission), or because different types of radiation affect those components in different ways (for example, ionization and displacement effects). In addition, some types of responses are application dependent, introducing yet another level of complexity in interpreting radiation responses.

Photonic devices can be used over a very wide range of wavelengths, and it is not possible to cover all aspects of optoelectronics in a course of this type. Many photonic devices are designed to be compatible with the three "windows" for fiber optics where $SiO_2$-based optical fibers allow light transmission over very long distance with low absorption, nominally 850, 1300 and 1500 nm, as well as with compatibility with commonly used detector materials. Those wavelength ranges are shown in Figure 1-1. The course briefly discusses photonic devices for the medium infrared region where it is generally necessary to use cooled detectors in order to obtain suitable signal-to-noise ratios. The far-infrared region -- above 5 $\mu$m -- is discussed only briefly in this course.

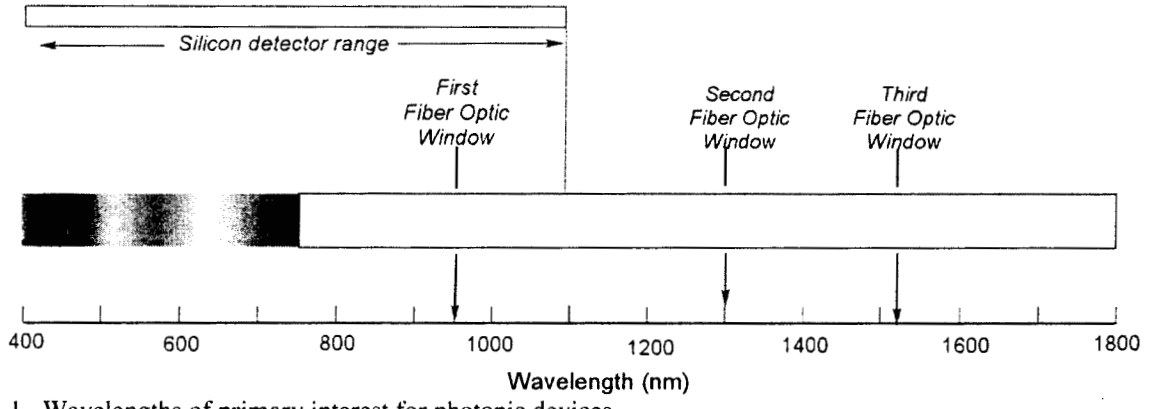



Figure 1-1. Wavelengths of primary interest for photonic devices.

This segment of the short course begins with a very brief review of radiation environments (Section 2), stressing the fact that we generally have a mix of high-energy electrons and protons in space environments, and that although radiation tests are often done with cobalt-60 gamma rays, those tests simulate only ionization damage, not displacement damage effects. Generally speaking, tests with gamma rays are inadequate to characterize the performance of optoelectronic devices in space. Energies for testing optoelectronic devices with protons and electrons are recommended, along with other important factors that affect the way that radiation testing and device evaluations are performed.

Sections 3-5 discuss fundamental effects in semiconductors that are important for photonic devices, including light emission and light absorption properties in direct and indirect semiconductors. The GaAs-AlGaAs system is used as an example of bandgap engineering to tailor devices towards specific wavelengths. Homostructure and heterostucture junctions are discussed, along with fundamental damage mechanisms.

The sixth and seventh sections discuss light emission from LEDs and laser diodes. The LED discussion contrasts three different LED technologies, and also discusses measurement and characterization techniques for

those classes of devices. Promising new laser technologies are included, particularly vertical cavity semiconductor lasers.

The eighth section discusses optical detectors. It includes conventional silicon photodiodes and phototransistors as well as III-V detectors. A brief discussion of infrared detectors is also included, as well as a discussion of noise in detectors and electronics.

The ninth section discusses the important topic of optocouplers, a good example of a device with multiple, complex failure modes. Optocoupler performance depends on physical characteristics of the emitter and detector and optical coupling compounds as well as on degradation of individual components.

The tenth and eleventh sections discuss other optoelectronic devices, including solar cells and charge-coupled devices. This is followed by another section with two examples of complex failure modes.

Section 13 discusses test techniques, including selection of proton energy, interpretation in the context of actual environments with a distribution of proton energies, and a limited discussion of SEE testing. The course is summarized in Section 14.

## *2 - Space Environments*

### *A. Overview*

Environments in space consist of protons and electrons that cause permanent damage effects in optoelectronic devices because there are very large numbers of them. For the most part, damage in optoelectronic devices is due to the large number of interactions that result from exposure to relatively high fluences of these particles. There are also galactic cosmic rays that are relatively few in number compared to protons and electrons. For galactic cosmic rays the primary concern is effects from the interaction of a *single* particle with the device (such as single-event upset). We are primarily concerned with permanent damage effects from protons and electrons in this part of the 2000 Short Course, except for a brief discussion of single-event effects testing at the end, and therefore very little attention will be given to heavy ions or transient effects from proton recoils.

Space environments are heavily influenced by trapped radiation belts around the earth (or trapped belts around other planets with magnetic poles) as well as by solar flares. This section is intended to provide only a brief summary of space radiation environments. A far more thorough treatment of radiation environments was given by J. L. Barth in Part 1 of the 1997 Short Course [Bart1], along with an earlier review paper by E. Stassinopoulous and J. Jaymond [Stas1]. Readers are encouraged to use those references for a more accurate and complete description of space environments.

### *B. Solar Flares*

Solar flares are important for many environments that are encountered by spacecraft. Solar flare activity is periodic, increasing during periods of intense sunspot activity on an eleven-year cycle. Most solar flares have very low intensity, but statistical results from the last three solar cycles have shown that there is a high probability of getting at least one solar flare with relatively high intensity during each solar cycle. Consequently, most spacecraft that operate within the 5-6 year period of more intense solar activity need to include an intense solar flare in their environmental requirements. However, basing the requirements on an extremely hard, intense flare (such as the flare that occurred in October, 1989) is overly conservative for most spacecraft because such intense flares occur very rarely.

Xapsos, et al. have developed a statistical model that incorporates data from the last 30 years to predict the maximum expected fluence from solar flares in a less conservative manner [Xap1]. Their model predicts upper bounds for various energies as shown in Table 1. Those values do not take geomagnetic and atmospheric shielding into account, but are applicable to geosynchronous or deep space environments. Table 1 shows the results of their model for various proton energies.

Table 1. Worst-Case Solar Event Fluences from Statistical Models [Xaps1]

| Energy (MeV) | Worst-Case Proton Fluence ($p/cm^2$) |
|---|---|
| 10 | $4.4 \times 10^{10}$ |
| 30 | $1.3 \times 10^{10}$ |
| 50 | $6.1 \times 10^{9}$ |
| 100 | $1.7 \times 10^{9}$ |

These values represent the total fluence from a single intense flare at the earth, not the total fluence from all solar flares that occur during an extended mission. However, the results show that the maximum expected fluence from a single flare is on the order of a few times $10^{10}$ $p/cm^2$. The fluence from solar flares also depends on the distance from the sun; the fluence falls off roughly as the inverse square of the distance. This is important for interplanetary missions.

### *C. Earth Orbiting Environments*

The trapped radiation belts that surround the Earth are the major contributors of radiation for low-Earth (LEO) and medium-Earth (MEO) orbits. Figure 2-1 shows a pictorial diagram of the proton belt and the outer electron belt (the inner electron belt is not shown).

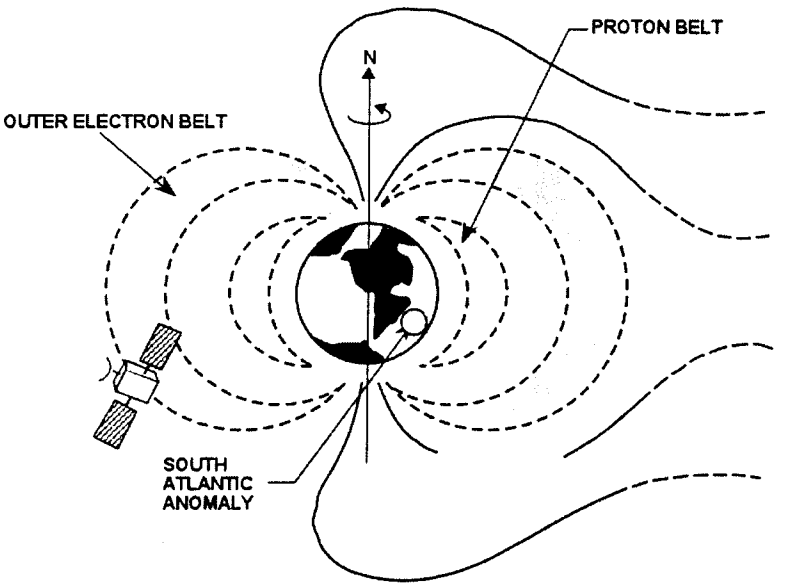



Figure 2-1. Pictorial diagram of the earth's radiation belts.

The proton environment that earth-orbiting spacecraft encounter depends on altitude and inclination. The intensity of the proton belts increases sharply with altitude, extending from about 800 to 15,000 km. The edges of the belt are not well defined, and the pictorial diagram of Figure 2-1 is somewhat misleading in that respect. Figure 2-2, after Stassinopoulous and Raymond [Stas1], shows contours of various proton energies for various altitudes along the equatorial plane (the belts are closer to the earth at higher latitudes). The figures also shows the L-shell parameter that is used to more accurately describe the position of the proton belts at other latitudes. Figure 2-2 shows that the edge of the proton belt actually begins at about 500 km, and that very high levels of proton radiation are encountered for altitudes above 2000 km. Most satellites that operate in low-earth orbits are constrained to altitudes below about 1500 km in order to avoid the intense radiation levels in the middle of the proton radiation belt.

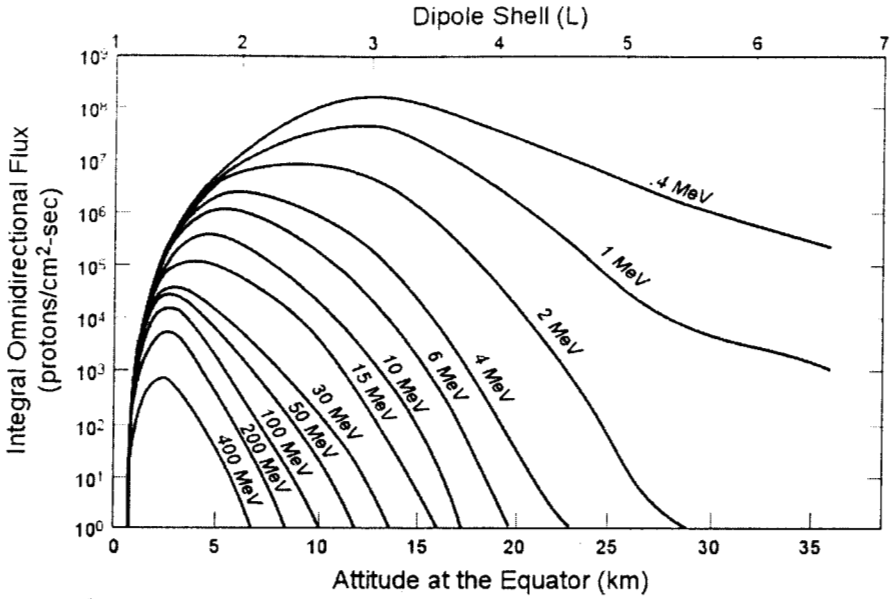



Figure 2-2. Energy contours for protons at various altitudes in the equatorial plane (after Stassinopoulos and Raymond [Stas1]).

For very low orbits (such as the Space Shuttle, approximately 305 km), the orbit is below the edge of the proton belt, and most of the protons occur when the spacecraft passes through the South Atlantic anomaly (SAA). An asymmetry in the proton belt causes it to extend (locally) to much lower altitudes in this region. The main reason for the SAA is that the earth's magnetic dipole is not exactly aligned with its rotational axis. Proton contributions from the SAA are also important at intermediate altitudes, but become less of a factor for spacecraft with high altitude and inclination. For high-inclination LEO and MEO orbits, protons from solar flares also make a significant contribution to the total proton fluence. However, at low inclination geomagnetic shielding keeps most of the solar flare particles out of the inner belt region, and the effect of solar flares can largely be ignored.

Spacecraft environments usually take shielding into account, typically assuming a spherical shell of aluminum that is a reasonable first approximation of the amount of shielding that surrounds most of the electronics. Moderate amounts of shielding reduce the number of electrons and low energy protons, but have little effect on protons with higher energy. Figure 2-3 shows how the internal environment of one orbit, the 705 km/98° orbit used for many of the missions to monitor climatic changes around the earth, is affected by shielding. In this case the spacecraft is assumed to operate for 5 years. With 100 mils of shielding, protons contribute more than twice as much total dose as electrons, and protons are even more dominant for thicker amounts of shielding. Although small amounts of shielding are effective in reducing the total dose in the "raw" environment, once the low energy particles are removed by the thin shield one is left with relatively energetic protons that become very difficult to shield unless very large additional amounts of shielding can be tolerated. Thus the effects of shielding are higher nonlinear. Note the presence of low levels of Bremsstrahlung radiation that place a lower limit on the effectiveness of shielding even for very thick shields.

Many different earth orbits are possible, and the number of protons expected during a mission depends on altitude, inclination and mission duration. Typical proton fluences (after moderate amounts of shielding) for LEO orbits range from about 2 x $10^8$/cm$^2$ for a short-duration space shuttle mission (305 km) to as much as $10^{11}$/cm$^2$ for a 1300 km polar mission operating for five years. Higher total dose levels are encountered for higher orbits or geostationary transfer orbits (see [Bart1] for details).

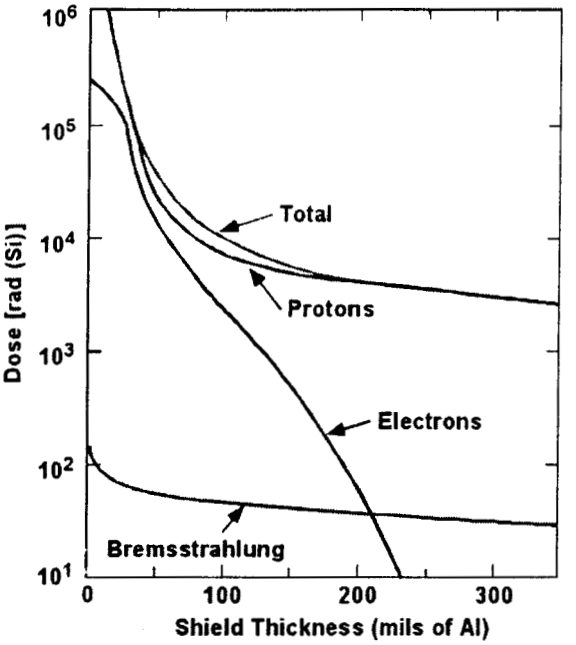



Figure 2-3. Effect of shielding on total dose from electrons and protons for a 705 km, 98° earth orbit.

Proton fluences are far lower for satellites that operate at high altitudes because they are beyond the edge of the proton radiation belt. For geostationary satellites (22.4 km), the proton belts do not contribute, and most of the proton fluence comes from solar flares.

### *D. Deep Space and Planetary Environments*

Proton levels in deep space environments are similar to those of geostationary orbits, with most of the contribution coming from solar flares (correcting, however for the falloff of solar flare intensity with distance from the sun). However, interplanetary missions may have to pass through intense radiation belts associated with other planets. Jupiter and Saturn both have trapped radiation belts that extend over far greater distances than the trapped radiation belts surrounding the earth. Planetary missions that need to traverse those regions can encounter very high levels of radiation. More details are given in Part 2 of the 1993 Short Course, presented by H. Garrett [Garr1].

Not all planets have magnetic poles. For example Mars has no trapped radiation belts, and consequently radiation levels near Mars are dominated by galactic cosmic rays and solar flares. The Mars atmosphere is much thinner than that of earth, but still provides sufficient shielding to reduce the radiation level on the Martian surface to relatively low levels.

## *3 - Charged Particle Interactions*

### *A. Basic Considerations*

When electrons and protons interact with materials the dominant energy loss process is by electromagnetic interaction with loosely bound electrons in the valence band. This process, know as ionization, elevates an electron from the valence band into the conduction band of the material, simultaneously creating an excess electron in the conduction band and a hole in the valence band (electron-hole pair). The process requires about 3.8 eV in silicon and 18 eV in silicon dioxide [McLe1], and thus only a very small amount of energy is absorbed by each ionization event. Ionization produces electron-hole pairs in insulators (such as silicon dioxide) as well as in semiconductors and the dominant damage effect from radiation in many semiconductor

devices is due to the way that the excess carriers generated by ionization are trapped at critical interface regions between the semiconductor and oxides. Examples include shifts in the threshold shift of gate and field oxides in MOS devices [Win1, Dres1], and gain degradation due to increase in surface recombination in many types of bipolar transistor structures [Peas1]. Ionization damage is so dominant for many important devices that system specifications often include only the net deposited dose from ionization in their specifications.

Another effect of ionization in insulators is the creation of color centers that affect the optical properties of the material. Color centers are created when an electron is trapped at an impurity site or vacancy in the material, or when the charge from ionization changes the valence state of an impurity atom. This effect is important in optical fibers, glass windows, and lenses. Absorption losses from color centers are wavelength dependent. In silicon dioxide absorption is much greater at short wavelengths. The degree to which color centers are created depends on the number and type of impurities within the material (for example, lead glass is extremely susceptible to darkening at relatively low levels of ionizing radiation). Some types of optical fibers are highly susceptible to formation of color centers (see Part 4 of the 1992 Short Course by E. J. Friebele [Frei1]). Color centers in quartz or other forms of pure silicon dioxide are generally only important for components with very long path length, and can usually be ignored for most space applications.

However, ionization is not the only energy loss mechanism. Electrons and protons can transfer energy to bound nuclei within the lattice via interactions with atomic nuclei. If the energy transferred during the collision is high enough the atom can be moved from its stable position, creating a disordered damaged region within the material. This process, referred to as displacement damage, is usually the dominant radiation effect for optoelectronic devices. The threshold energy to create a displaced atom in silicon is about 13 eV, but the total energy transferred to the lattice can exceed 10,000 eV for complex damage cascades [Srou1]. On the average, each displacement event deposits much more energy than ionization events. Ionization loss processes still dominate from the standpoint of overall energy loss by the incident particle because the cross section for nuclear interactions (displacement) is about a factor of $10^4$ lower than that due to ionization due to the small size of the nucleus.

We will concentrate most of the discussion in this section of the course on displacement damage in various kinds of optoelectronic devices because it turns out to be the most important damage mechanism for most photonic devices. In addition to effects on semiconductor properties displacement damage can alter basic optical properties of materials (such as the index of refraction), but such effects are usually only important for very high fluences that are above those encountered in typical spacecraft and will not be considered further in this course.

Single-event effects from galactic cosmic rays or proton-induced reaction are also important for many photonic devices, and are addressed in Part 1 of this year's short course. Although single-event effects are generally omitted from this part (Part 2), some aspects of single-event radiation testing of optoelectronic devices are discussed in the next-to-last section.

### *B. High-Energy Protons*

Proton displacement damage depends on energy, and the energy dependence must be taken into account in order to relate laboratory test results -- usually done at a single proton energy -- to equivalent effects on devices when they are exposed to real space environments with a distribution of proton energies. The energy dependence depends on the material as well as the specific property that affects the optoelectronic devices that are being considered, i.e., lifetime damage, mobility degradation, or carrier removal.

During the last 15 years considerable progress has been made in calculating the energy dependence of the component of energy loss that goes into displacement damage. The term non-ionizing energy loss (NIEL) is used to describe this calculation. Figure 3-1 shows the energy dependence of NIEL for protons in silicon as derived by Summers, et al. [Sum1] and experimentally verified with measurements of lifetime damage in discrete transistors. The damage increases rapidly at low energies (with approximately a 1/E dependence) basically because it takes longer for low-energy protons to traverse the target material, providing a longer time for the interaction to occur. Although this is effectively a collision between the proton and the nucleus, the interaction is actually electromagnetic at lower energies which is the reason for the increase in NIEL at low

energies. The NIEL value for 1-MeV equivalent neutrons is also shown (for neutrons the interaction is kinematic rather than electromagnetic). Protons with energy of 200 MeV have about the same NIEL value as 1-MeV equivalent neutrons, which is of interest because of the body of data that exists for neutron degradation.

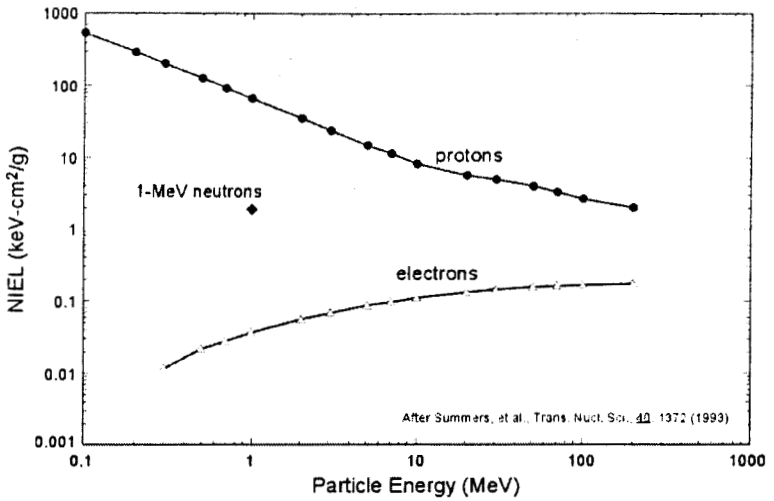



Figure 3-1. Dependence of NIEL on proton and electron energy for silicon.

Similar calculations have been done for NIEL in GaAs, but the results are less consistent with experimental results. Figure 3-2 shows the calculated energy dependence [Summ2, Summ3] along with experimental results obtained by Barry, et al. for light-emitting diodes [Barr1]. The calculated values of NIEL agree reasonably well with older experimental data for JFETs, which change resistivity after irradiation because of carrier removal. However, the calculated values disagree at high energies compared to the more recent experimental results for LEDs, which degrade because of changes in minority carrier lifetime. The discrepancy between the NIEL calculations and LED experimental results is about a factor of three at 200 MeV. Note that there is reasonable agreement at energies below approximately 80 MeV.

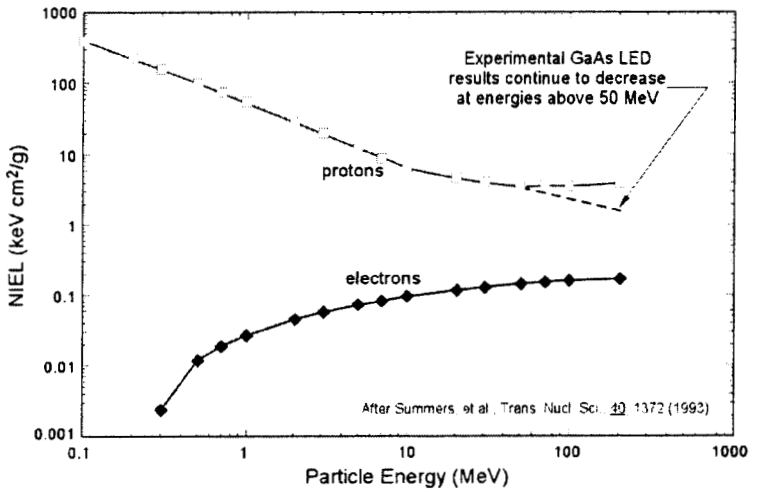



Figure 3-2. Dependence of NIEL on proton and electron energy for GaAs.

For space environments the increased effective NIEL at low energies makes the low energy component of the proton energy spectrum relatively more important from the standpoint of damage in devices. Figure 3-3 shows the "raw" proton spectrum for a high-inclination Earth orbit [New 1] along with *adjusted* spectra for silicon and GaAs displacement damage that take the energy dependence into account by weighting each energy interval by the relative displacement damage effect. For GaAs, the effect is to approximately double the raw spectrum (system specifications often include the total number of protons as part of their requirements). There is less difference between the raw and adjusted spectra for silicon. These equivalence factors depend on the amount of shielding that is present, along with the initial spectrum. As the amount of shielding is increased, the mean energy becomes higher and the difference between the raw and adjusted spectrum becomes smaller.

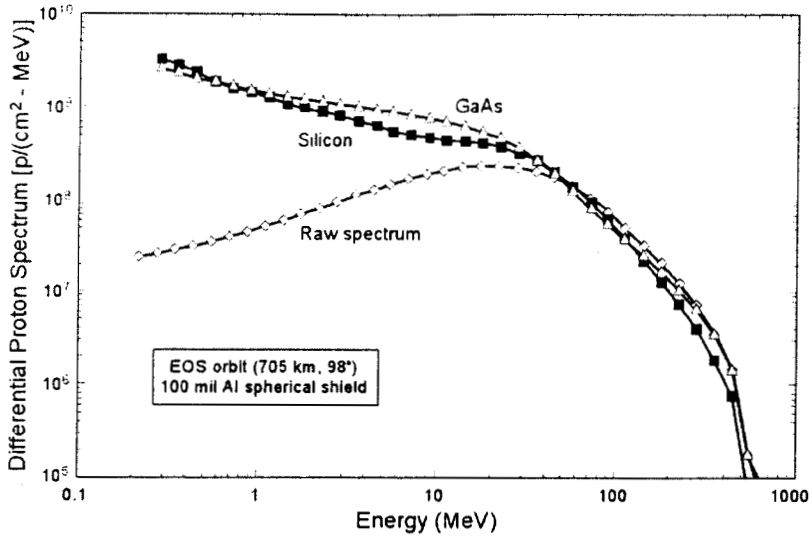



Figure 3-3. Typical energy spectrum; example of "adjusted" spectrum and difference between "raw" and "effective" fluences.

Proton testing is usually done at only a single energy (except for special studies to investigate energy dependence) because it is too costly to include tests at several different energies on a routine basis. Although in principle any energy can be used (subject to range limitations), the best practice is to select an energy for testing that is near the peak in the differential energy spectrum, approximately 50 MeV for many systems (see Figure 3-3). If the energy is too low, the range of the protons may be insufficient to penetrate the package and dead layer of the device, leading to uncertainties in the energy of the protons that actually penetrate the active region. Solar cell data is usually based on 10 MeV protons because the cells are generally shielded only by very thin cover slides, but higher energies are recommended for testing most other optoelectronic components because there is much more extraneous material in lenses, packaging and subsystem enclosures, leading to higher mean energies in the spectrum of protons that actually reaches the device.

Using energies above 80 MeV is potentially a problem for III-V devices because of the discrepancy in NIEL between theoretical and experimental work discussed earlier. When high energies are used there is no ambiguity about the measured results, but the interpretation of how the damage affects devices with a continuous proton spectrum then depends on the assumed energy dependence for NIEL. If the NIEL calculations are revised at a later date (or inaccurate), then the interpretation of the effect of the actual protons in the environment will be incorrect. For example, 200 MeV protons are readily available at one high-energy proton facility, and have frequently been to test optoelectronic devices. If the NIEL values for liftetime damage are used, the damage for a typical spectrum of protons will be *a factor of three higher* than if the NIEL values for carrier removal (or the theoretical calculations of NIEL) are used. Testing with 50 MeV protons reduces such uncertainties to about 20-30%.

### *C. High-Energy Electrons*

Electrons also produce displacement damage, with an energy dependence that is quite different than for protons. As shown previously in Figures 3-1 and 3-2, NIEL for electrons increases with energy, in contrast to protons. With electrons there is a sharp energy threshold -- approximately 150 keV for silicon -- below which displacement damage does not occur [New2]. For GaAs the threshold energy is about 250 keV. Although electron displacement damage is potentially important, the NIEL values are two to three orders of magnitude lower than for protons. This, in combination with the reduction in electron fluence from moderate amounts of shielding usually makes electron displacement damage less important than proton displacement damage in earth-orbiting environment. Exceptions are cases where there is very little shielding (such as for solar cells), or interplanetary missions where the electron spectrum may be much harder than the spectrum of electrons in the earth's radiation belts. For example, electrons in the Jovian radiation belts have energies up to 500 MeV.

## D. *Displacement Damage Defects*

Displacement damage is a complex process, and the nature of the defects that are created depends on the energy that is actually transferred to the lattice site, which varies over a broad range. The displacement energy threshold is about 9 eV for GaAs and 13 eV for silicon. Displacement events that transfer relatively small amounts of energy create isolated vacancy-interstitial pairs along the particle track (Frenkel pairs). When larger amounts of energy are transferred the nature of the localized damage sites changes. In silicon damage clusters tend to form over distances of about 60 Å. The clusters are charged by the energy deposition process (the displaced atom at the site of the original collision creates the clustered damage site by electromagnetic interaction with other atoms in the lattice as it dissipates energy). The cluster damage sites are not completely stable, and some of the damage gradually anneals with time. Because the cluster is charged, annealing occurs much more rapidly when current passes through the region [Greg1]. This is termed injection-enhanced annealing.

One would expect that when a lattice atom absorbs energy that is several orders of magnitude above the threshold energy for displacement that the result would be a very large localized damage region, spreading radially from the collision site. However, this is not the case. Figure 3-4 shows an example of how the damage progresses through the lattice, based on a computer model [Lint1]. In this case the initial recoil atom absorbs 60 keV from the collision. The effect of the higher recoil energy is to create a cascade of damaged cluster regions, each with about the same dimension as the clusters that are created by events that transfer more moderate amounts of energy to the lattice. Frenkel pairs are also produced along the track of the recoil. In this example clusters are produced nearly 1000 Å beyond the site of the initial collision.

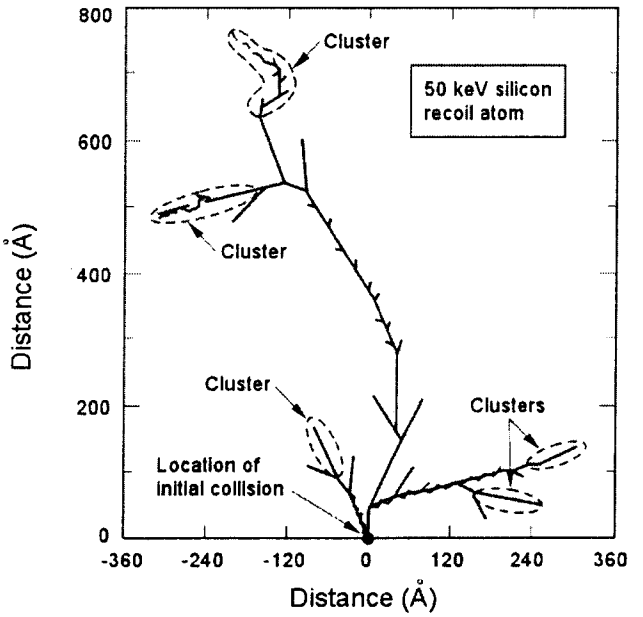



Figure 3-4. Damage cascade structure, calculated from a theoretical model. Note that there are several different damaged regions along the track with about the same disordered extent, not a single region with large defect density.

Figure 3-4 illustrates the different character of the microscopic defects that are created by displacement damage. Based on these considerations, it is rather surprising that the concept of a general equivalence for displacement damage will work at all. Several aspects of damage are not treated directly by NIEL calculations, including the issue of damage stability and annealing. As noted by Summers [Sum1, Short Course] as well as by C. Marshall and P. Marshall in Part 3 of the 1999 Short Course [Mars1], it is surprising that the NIEL results agree as well as they do considering that they are applied to different types of particles over a wide range of energies. Thus, it is important not to expect too much from NIEL calculations. The concept of NIEL is extremely valuable when making damage comparisons, but there are differences in the damage mechanisms that are likely to lead to discrepancies if the concept is pushed too far. Table 4-1, after Dale, et al. [Dale1] shows the way that defects are distributed between clusters and isolated point defects for electrons and protons in silicon for different amounts of deposited energy.

Table 3-1. Partition of Microscopic Damaged Regions

| Energy Interval Range of Primary Knock-on Atoms | Defect Structure | Percent of Defects | | | |
|---|---|---|---|---|---|
| | | Electrons | | Protons | |
| | | 4.1 MeV | 53 MeV | 8.3 MeV | 60 MeV |
| 25 to 1000 eV | Point defects | 97 | 44 | 34 | 12 |
| 1000 to 10,000 eV | Single cascade and isolated point defects | 3 | 28 | 21 | 7 |
| Above 10,000 eV | Several subcascades and isolated point defects | 0 | 28 | 45 | 81 |

## *E. Displacement Damage Effects in Semiconductors*

Displacement damage affects various properties of semiconductors. The importance of displacement effects and their interpretation is different for various optoelectronic devices. These effect are discussed below; it is necessary to have a reasonably good understanding of the underlying structure and principles of the operation of devices in order to put these effects in the proper context. Those issues are discussed in more detail in later sections of this part of the Short Course.

Lifetime Damage

The semiconductor property that is most sensitive to displacement damage effects is minority carrier lifetime. Lifetime damage can be described by the equation

$$\frac{1}{\tau} - \frac{1}{\tau_0} = \frac{\Phi}{k}$$

where $\tau_0$ is the initial lifetime, $\tau$ is the lifetime after irradiation, $\Phi$ is the particle fluence, and K is the damage constant. The relationship between the reciprocal liftetime and the reciprocal of the damage constant is linear over a wide range of fluence values, until second-order effects (such as carrier removal) become important. The damage constant differs for n- and p-type material, and also depends on doping level. This equation is widely used to describe lifetime damage, but does not address damage stability or annealing. It can be applied to other device parameters that are sensitive to minority carrier lifetime, such as common-emitter current gain [Mess1].

Figure 3-5 shows how protons affect minority carrier lifetime in n-type silicon for 50-MeV protons. In this example the silicon doping concentration is $10^{16}$ atoms/cm$^3$. Devices that depend on long carrier lifetime for operation are degraded significantly at levels between $10^{10}$ and $10^{11}$ p/cm$^2$, but devices that begin with shorter initial lifetimes (or are insensitive to lifetime changes, such as MOS transistors) will be unaffected until much higher levels of radiation. Thus, the lower bound for concern about proton displacement effects is on the order of $10^{10}$ to $10^{11}$ p/cm$^2$, for 50-MeV protons.

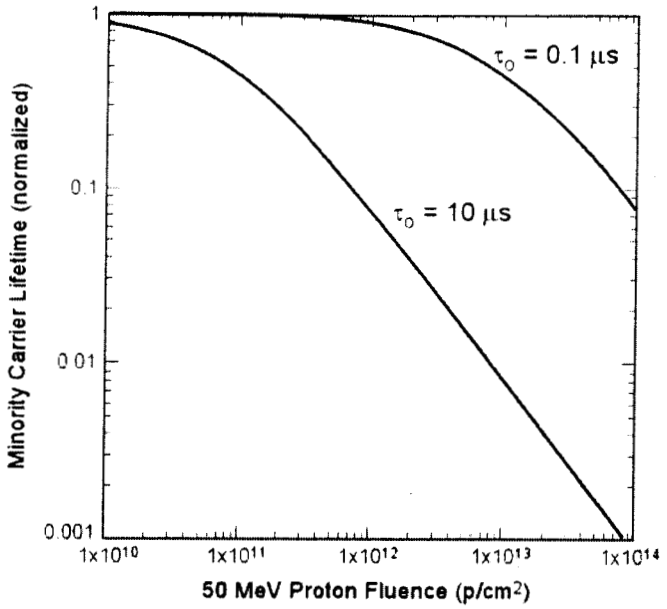



Figure 3-5. Effect of 50-MeV protons on minority carrier lifetime in silicon.

Similar results hold for GaAs and related materials, but the damage constants are different. The operation of many GaAs semiconductors (e.g., JFETs and MESFETs) does not depend directly on carrier lifetime, and for this reason relatively little data exists in the literature on lifetime damage in GaAs and other III-V materials. Data on some types of LEDs are an important exception, but in most LED studies the lifetime has not been measured, just the overall performance of the LED after irradiation.

Carrier Removal

Carrier removal occurs when radiation introduces deep impurity levels that can act as amphoteric impurities. Carrier removal is the result of compensation of the initial dopant atoms by the radiation-induced impurities. The carrier removal rate depends on doping level as well as whether the material is n- or p-type. The units are $cm^{-1}$. Table 3-2 below shows approximate carrier removal rates for electrons and protons in silicon and GaAs as a rough guide to the magnitude of carrier removal effects. For material with a doping concentration of $10^{15}$ atoms/$cm^3$ carrier removal will begin to become significant at proton fluences on the order of $10^{12}$ p/$cm^2$, about two orders of magnitude higher than the "lower bound" fluences that are of concern for lifetime degradation. Higher fluences are required in order for carrier removal to be important in materials that are doped at higher levels.

Table 3-2. Approximate Carrier Removal Rates for Protons and Electrons

| Particle | Silicon | GaAs |
|---|---|---|
| 1 MeV electrons | 0.15/cm | 0.6/cm |
| 50 MeV protons | 12/cm | 30/cm |

For small numbers of defects the deep-level impurities that are introduced by radiation act as amphoteric dopants. However, the impurities are not located exactly at the center of the bandgap and consequently the net doping concentration will shift towards n- or p-type when very large numbers of defects are present. For example, n-type silicon becomes p-type after high levels of neutron or proton irradiation[[Leme1]. This will be discussed further in Section 11.

Mobility

Carrier mobility is also affected by displacement damage. Radiation-induced defects increase carrier scattering, reducing carrier mobility. Mobility degradation is important for MOS devices because mobility is directly related to transconductance, but this generally requires much higher particle fluences than the levels required to produce large changes in minority carrier liftetime. Lifetime damage and carrier removal are usually the most important material parameters for optoelectronic devices, and mobility degradation will not be considered further in this course. Part X of the 19XX Short Course discusses mobility in more detail [Mob1].

## *4 - Key Properties of Semiconductors Used in Optoelectronic Devices*

### *A. Band Structure*

Basic features of the interaction of light quanta (photons) with semiconductors can be explained with the band theory of solids, which describes the energy and momentum of holes and electrons within a semiconductor crystal (a periodic lattice). Allowable quantum states within a periodic structure are described in terms of crystal momentum, **p**, and wavevector, **k**, corresponding to the electron (or hole) states within the lattice. Note that **p** and **k** are both vectors; they are "natural" descriptions of solutions to the quantum-mechanical problem of stable states within an idealized periodic crystal (Bloch functions). As shown in reference [New4], the wave vector **k** is related to lattice periodicity and is a basic characteristic of eigenvectors that are allowable quantum states within the lattice. Figure 4-1 shows the way that energy depends on the wave vector for a simple periodic lattice. Near the boundary of the potential change (effectively a lattice site) there is a region of forbidden energy that arises from the quantum properties of the crystal. This leads to an energy gap between different wave vector states, as shown by the periodic discontinuities in the figure.

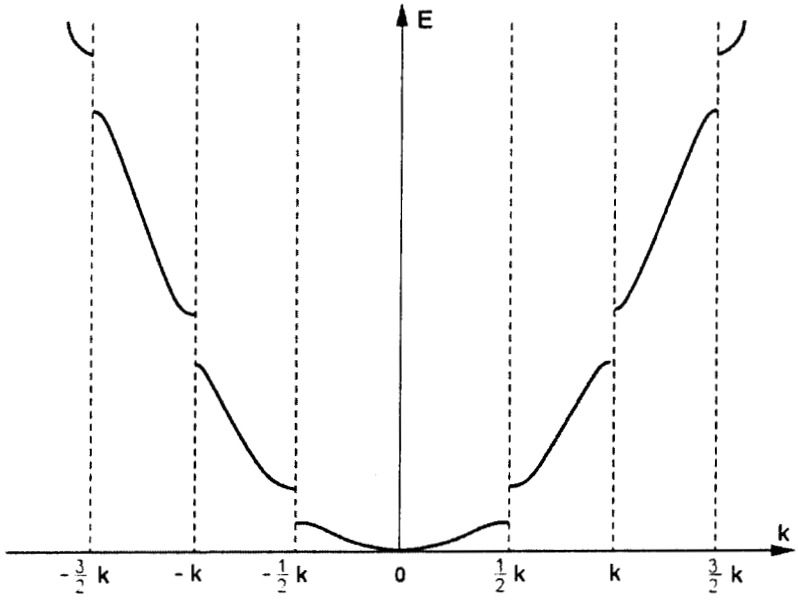



Figure 4-1. Relationship between energy and wave vector for a simple periodic potential showing energy gaps between allowed wave vector states.

With this formalism, the momentum of a hole or electron within the crystal is related to the wave vector that describes it by

$$\mathbf{p} = h\mathbf{k}/2\pi \qquad (1)$$

where h is Planck's constant ($6.62 \times 10^{-27}$ erg-s). The momentum generally has a different relationship with **k** for holes and electrons. Note that the crystal momentum for either electrons or holes differs from the classical momentum of a free particle because the particle behaves differently under the influence of the crystal

potential. This difference can be taken into account by using the concept of an effective mass, related to the curvature of the energy band structure. The band structure of real materials is far more complex than that of the simple model described by the E-k relationship in Figure 4-1 (see [Sze1] for more detail).

For optoelectronic devices the key feature of the band structure is the location of the minimum energy transition point in k-space. Figure 4-2 shows examples of two different band structures. For the "direct" semiconductor, the minimum energy in the conduction band occurs at the same k-value as the maximum energy in the valence band. This allows transitions between the two energy bands for the same value of **k** that can be accomplished by absorption or emission of a single photon. GaAs and several other types of semiconductors have band structures that allow this type of direct transition.

Figure 4-2(b) shows a different structure where the minimum conduction band energy occurs for a different value of **k** than for the k-value corresponding to the maximum conduction band energy. This type of indirect transition requires additional energy, typically provided by a phonon from the crystal lattice, to provide the additional momentum that is required to change the wave vector, **k**. Although direct transitions can still occur, they have very low probability, and require more energy than indirect transitions. Silicon is an example of a material with a band structure that is dominated by indirect transitions. Materials with indirect bandgap structures are very inefficient at producing light, which prevents their use in LEDs and lasers.

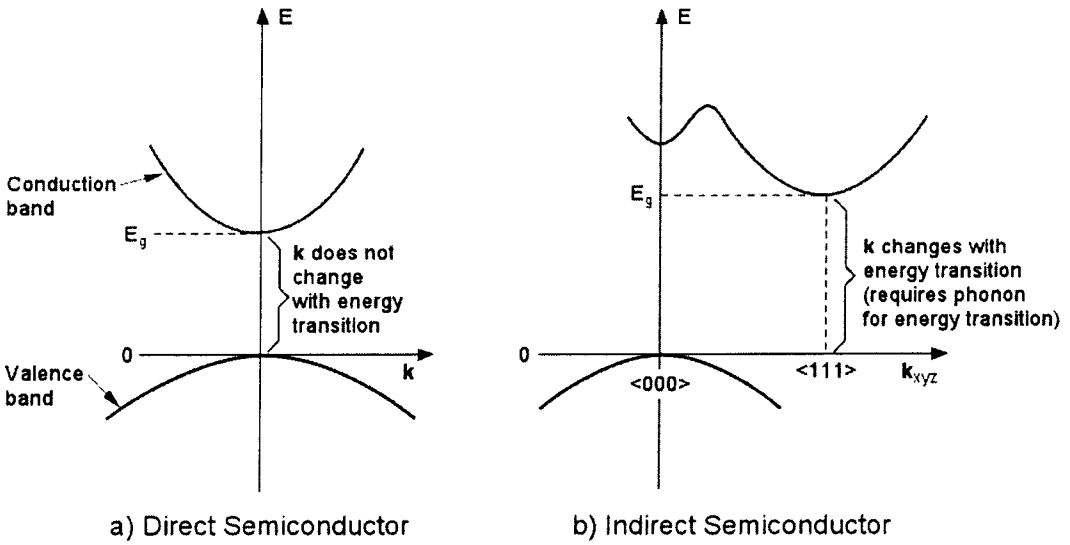



Figure 4-2. Band structure of direct and indirect semiconductors showing allowed wave vector transitions.

The examples in Figure 3-2 apply for crystals with low impurity concentrations where the band edges are well defined. For high impurity concentrations (above $10^{18}$ $cm^{-3}$), the band edges are "smeared" because of the presence of many different impurity atoms. The result is a lowering of the bandgap for high impurity concentrations. The presence of such bandtails affects emission and absorption properties because transitions can then occur at energies that are below the bandgap energy for the basic material when it is lightly doped.

### *B. Optical Emission and Absorption*

Although many properties of light are the result of its behavior as a wave, absorption and emission in atoms and solids requires that light be viewed from the standpoint of quantized photons (a beam or pulse of light corresponds to a packet of photons). The energy of a photon with frequency $\nu$ and wavelength $\lambda = c/\nu$ (c is the velocity of light) is given by the relationship

$$E = hc/\lambda \qquad (2)$$

The practical formula $E = 1.24 / \lambda$ is useful to remember, where E has units of electron-volts, and $\lambda$ has units of $\mu m$.

The simplest optical absorption process in a solid corresponds to direct absorption by a particle in the valence band that causes its energy to increase to (or beyond) the energy of the valence band. When this occurs, two particles are effectively created: an electron in the conduction band, and a hole in the valence band. In order for this process to occur the photon energy must exceed the band gap in the material. The corresponding process for photon emission occurs when an electron in the conduction band loses energy by emitting a photon, and falls into the valence band. A hole and electron are "destroyed" or annihilated when photon emission occurs.

As discussed earlier, the band structure of a semiconductor is important in determining whether direct absorption of a photon is the dominant process for transitions between states. Although direct transitions are possible in materials with indirect bandgaps, the probability of direct transitions in indirect semiconductors is small. Instead, the dominant process involves a second particle (usually a phonon of energy $\hbar\varpi$ that is present because of thermal motion of atoms within the crystal lattice).

Because of the additional energy provided by the phonon, the energy associated with the transition is slightly different from the bandgap energy for an indirect semiconductor. Indirect transitions take much longer to occur than direct transitions, and cause the dominant recombination process to be strongly affected by nonradiative recombination centers in a typical semiconductor junction instead of the radiation recombination that dominates for direct materials. This effectively means that indirect semiconductors cannot be used for laser diodes or LEDs. Other transitions are possible besides direct transitions, but will not be discussed here.

### *C. Absorption Coefficient*

Absorption of light can be described from a macroscopic point of view by the equation

$$I = I_0 e^{-\alpha x} \tag{4}$$

where $I_0$ is the initial light intensity at the surface, I is the reduced light intensity at a distance x within the material, and $\alpha$ is defined as the absorption coefficient. The quantity $1/\alpha$ (absorption depth) corresponds to the depth in the material where the intensity has fallen to 1/e times its initial value.

The absorption coefficient depends on wavelength. The dependence is abrupt for semiconductors with direct bandgap, but it is much more gradual for indirect semiconductors such as silicon. Figure 4-3 shows the absorption coefficient of silicon, GaAs, and InGaAs [New5]. The absorption coefficient falls to very low values once the energy of the photon is below the energy corresponding to the bandgap in the material.

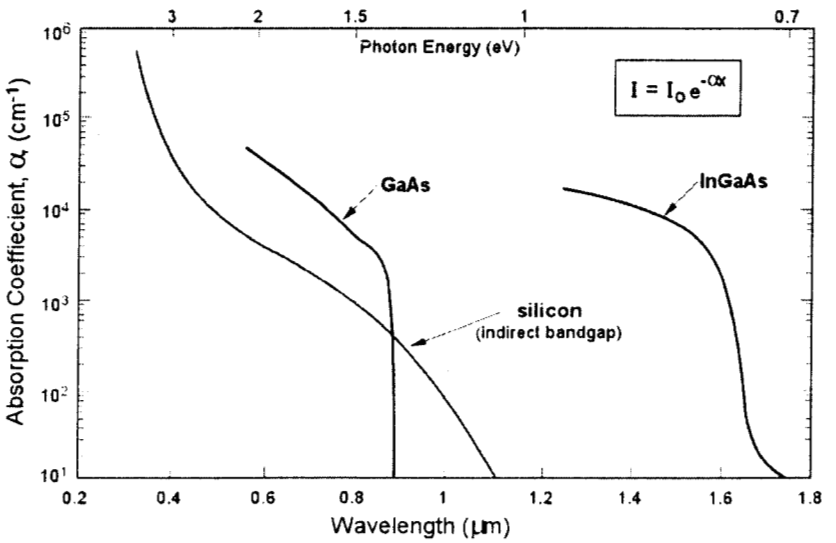



Figure 4-3. Absorption coefficients of GaAs, InGaAs and silicon. Note the more gradual dependence of $\alpha$ on wavelength for silicon, which has an indirect bandgap.

The absorption coefficient depends on temperature primarily because the bandgap is temperature dependent [Bgt1]. For silicon, the absorption coefficient changes about -1%/°C for temperatures near room temperature. Although this may appear to be a small dependence, it causes substantial differences in light absorption even over a moderate temperature range. Note that because $\alpha$ depends on wavelength, the absorption depth also depends on wavelength. The absorption depth is also affected by temperature. For photodetectors these factors cause the responsivity to change with wavelength. This is discussed in more detail in Section 8.

### *D. Quantum Efficiency*

When a photon is absorbed within a semiconductor, equal number of holes and electrons are created. For a steady-state process, this results in an excess carrier density. The excess carrier population can decrease either through radiative or non-radiative processes. Typical non-radiative processes include lattice vibrations (thermal energy) or recombination through impurity centers. The internal quantum efficiency is defined as the fractional number of non-radiative processes relative to radiative processes. Carrier lifetimes can be assigned to radiative and non-radiative processes, which allows the internal quantum efficiency $\eta$ to be defined as

$$\eta = (1 + \tau_{rad}/\tau_{nr})^{-1} \tag{5}$$

where $\tau_{rad}$ is the radiative lifetime, and $\tau_{nr}$ is the non-radiative lifetime. In order to keep the quantum efficiency high, the ratio $\tau_{rad}/\tau_{nr}$ should be as small as possible.

The internal quantum efficiency places an upper bound on the efficiency of absorption and radiative processes, but the net efficiency of an optical emitter or detector must also consider other processes that lower the effective efficiency when one considers the overall operation of the device. The external quantum efficiency includes additional factors such as absorption of light in highly doped or transition regions, and reflection from interfaces which limit light transmission.

### *E. Snell's Law*

Light is refracted when it passes through two surfaces of different refractive index. A light wave incident at an angle $\Theta_1$ in a material with refractive index $n_1$ will emerge at an angle $\Theta_2$ after it passes through the interface. The two angles depend on the refractive indices of the two media as described by Snell's law

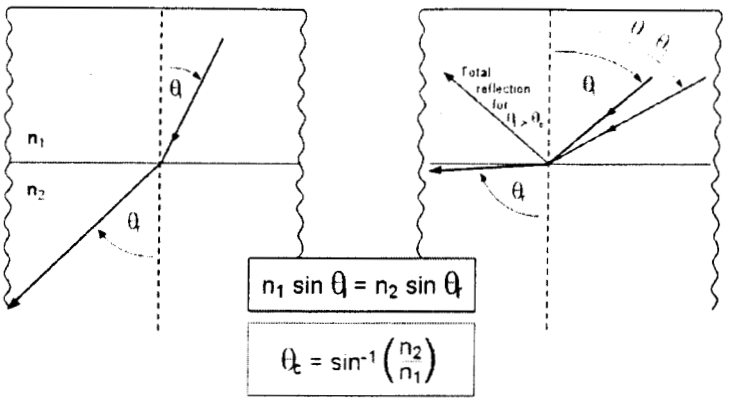
$$\frac{n_1}{n_2} = \frac{\sin\Theta_2}{\sin\Theta_1} \tag{6}$$

where $n_1$ and $n_2$ are the refractive indices of the two materials. For the case where the second material has a *higher* index of refraction than the first, this equation can be satisfied for all angles.

However, for the case where the second material has a *lower* refractive index than the first material, this equation can no longer be satisfied when the angle of the refracted wave exceeds 90 °. Once the angle of incidence exceeds the critical angle defined by

$$\Theta_c = \sin^{-1}(n_1 / n_2) \tag{7}$$

the incident wave can no longer be transmitted through the lower index medium. Instead, it is reflected at the interface (total internal reflection), as shown in Figure 4-4. The angle $\Theta_c$ is known as Brewster's angle. Total internal reflection is extremely important in determining how optical power within LEDs and laser diodes can be transmitted externally because the high refractive index of III-V compounds causes the Brewster angle to be low, about 16 degrees. This means that light that is incident at the surface at angles above 16 degrees will be totally reflected, and cannot be extracted from the material.




a) Moderate Angles b) Large Angles

Figure 4-4. Refraction at an interface where $n_2 < n_1$.

It is also the principle by which light can be guided through extremely long distances in optical fibers, which are coated with a cladding layer of lower refractive index than the fiber core. Table 4-1 shows the refractive index for several semiconductor materials as well as glass.

Table 4-1. Index of Refraction for Various Materials

| Type | Refractive Index |
|---|---|
| Si | 3.45 |
| AlAs | 2.97 |
| GaAs | 3.59 |
| $SiO_2$ | 1.46 |
| air | 1.00 |

The index of refraction depends on wavelength for all materials. For semiconductors, the doping level also affects n; it decreases with increasing doping density. For example, high-purity n-type GaAs has an index of refraction at 0.9 $\mu$m of 3.54. The index of refraction is lowered to 3.24 -- a 10% decrease -- for n-GaAs that is doped with a concentration of $6 \times 10^{17}$ $cm^{-3}$.

## *F. Optical Transmission*

Optical transmission through an interface of two materials depends on the refractive index. At normal incidence, optical transmission is given by the relationship

$$\rho = \frac{n_o - n_1}{n_o + n_1} \qquad (9)$$

where $\rho$ is the ratio of the amplitudes of the incident and reflected waves that describe the light path, and $n_o$ and $n_1$ are the refractive indices of the first and second materials, respectively. Note that optical power depends on $\rho^2$. If $\rho$ is negative, the reflected wave is opposite in phase to that of the incident wave. The above equation is a reasonable approximation for moderate angles of incidence. However, at more extreme angles the amplitude and phase of the reflected wave depend on angle in a complex manner (see Reference Z).

For the case where the first material is air and the second material is glass, $n_1 = 1.5$, and the reflected optical power is 0.04; that is about 4% of the incident energy is reflected from the interface. For GaAs, $n_1$ is 3.54 and about 33% of the incident energy is reflected from an air-GaAs surface.

Note that total internal reflection limits the angle over which light can be transmitted for the case where the light wave originates within a material with high index of refraction. The net amount of light that is transmitted depends on the compound effects of the limitations imposed by the Brewster angle and Fresnel losses. This is shown in Figure 4-6 for GaAs and the more frequently encountered condition of a glass-air interface. For GaAs, only about 2% of the internal light is transmitted from GaAs to air compared to silicon.

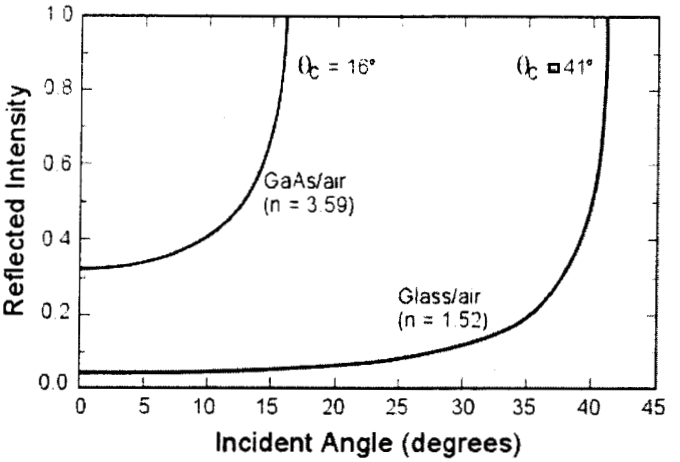



Figure 4-6. Effects of Fresnel loss and total internal reflection on light transmission for GaAs and silicon-dioxide with an air interface.

The use of an optical coating or *index matching* medium can increase the amount of energy that is transmitted through such interfaces. Nearly all optoelectronic devices use coatings or matching compounds to increase optical transmission at interfaces. Index matching coatings are important in radiation testing because their transmission properties may be affected by ionizing radiation.

## *5 - Properties of III-V Semiconductors*

### *A. Electrical Properties of Basic Materials*

Many III-V semiconductors are direct-bandgap materials, and therefore can be used to fabricate light-emitting diodes, laser diodes, and other optoelectronic devices that are not possible with silicon or other materials with indirect bandgaps. Gallium arsenide is an example of a typical III-V direct-bandgap semiconductor, but there are several others of interest, as shown in Table 4-1. The bandgap energy and wavelength corresponding to the absorption edge are shown in the table. Light is not absorbed beyond the absorption edge that corresponds to the bandgap energy, which provides an upper limit to the material.

Table 5-1.
Bandgap Energies of Some Direct-Bandgap Semiconductors

| Material | Bandgap Energy (eV) | Absorption Edge ($\mu$m) | Typical Application |
|---|---|---|---|
| GaAs | 1.42 | 0.87 | LED/laser technology |
| InP | 1.35 | 0.92 | LED/laser technology |
| GaSb | 0.72 | 1.72 | Long wavelength LEDs |
| InAs | 0.36 | 3.44 | Long wavelength detectors |

It is much more difficult to grow high-purity III-V devices compared to semiconductors with a single element (such as silicon and germanium) because of the difficulty of maintaining constant growth of both types of atoms in the crystal. A great deal of effort has been spent in developing advanced methods of growing single-crystal devices in such materials. The presence of two different types of atoms in the III-V lattice makes it possible to form admixtures of different constituents, changing the bandgap. Although this is a problem when attempting to grow pure GaAs, this feature of III-V devices can be used to advantage because the properties of the semiconductor change when the ratio of the constituents is altered. This makes it possible to tailor the wavelength of light-emitting diodes and laser diodes by varying the relative composition. Only two such material classes will be discussed in these notes; others are discussed in references XXX.

### *B. Ternary and Quaternary Materials with Variable Bandgap*

Although one usually thinks of III-V semiconductors as having only two constituents, it is possible to form solid solutions that incorporate a third type of atom and still retain the single-crystal characteristics of a semiconductor. For example, by introducing varying amounts of aluminum, a solid solution of aluminum-gallium-arsenide is formed. The bandgap of this solid solution can be adjusted over a relatively wide range by varying the amount of aluminum [Alg1]. The resulting composition is designated $Al_xGa_{1-x}As$, where the subscript x denotes the fractional concentration of aluminum. Figure 5-1 shows how the bandgap and nominal wavelength depend on the fractional aluminum concentration. Note, however that the band structure changes to that of an indirect semiconductor for aluminum concentrations above 30%.

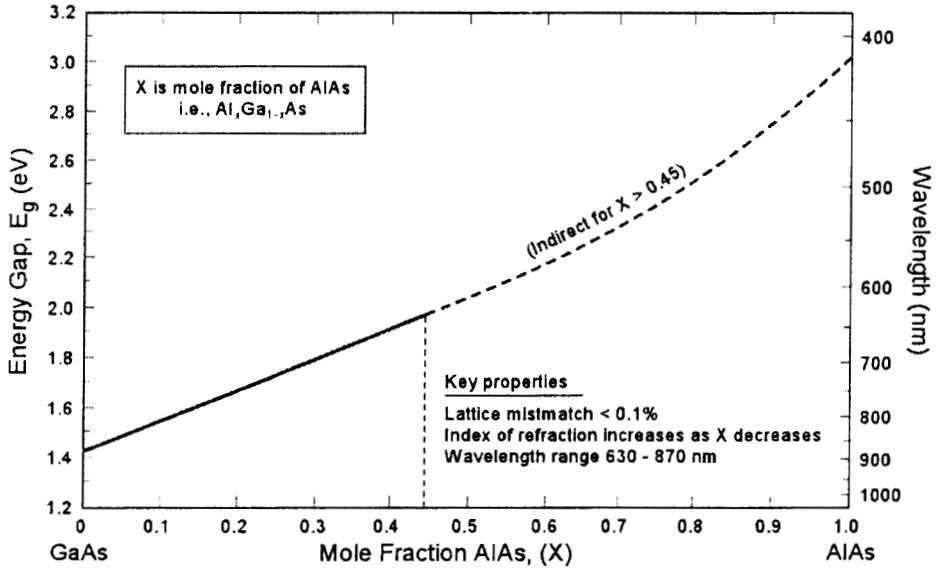



Figure 5-1. Bandgap and wavelength of the AlGaAs system

AlGaAs can be tailored to operate over a wavelength range of approximately 630 to 870 nm. The index of refraction also changes with composition, increasing as the amount of aluminum increases. One unusual property of AlGaAs is that the lattice constant changes only slightly as the composition is varied, making it possible to grow crystals with low defect concentrations.

Other ternary semiconductor systems are possible besides AlGaAs, with different wavelength ranges, but generally have lattice constants that are not closely matched to other material types except for specific compositional ratios. That restricts their use to specific wavelengths where adequate lattice matching occurs. To overcome this basic material limitation, quaternary compounds have been investigated for LEDs and lasers that operate at longer wavelengths with better lattice matching than most ternary materials. AlGaAsP is widely used because of excellent lattice matching to fabricate devices with wavelengths of about 1.3 μm [Ref Y].

## C. *Heterojunctions*

A heterojunction is formed when two semiconductors that have different bandgap energies are joined together. If the two materials are of different types (e.g., n-and p-type semiconductors), then the junction formed by the two materials functions much like the p-n junction of a conventional semiconductor material (homojunction), providing a barrier to minority carrier injection with rectifying properties [Het1]. If the two materials in the heterojunction are of the same type, there is no rectifying junction but the heterojunction helps to confine minority carriers within the region with lower bandgap. In addition to the electrical properties, the two materials in a heterojunction have different refractive indices which is an important feature for LEDs and laser diodes. Figure 5-2 shows the bandgap structure of a GaAs/AlGaAs heterojunction. The GaAs region is p-doped with a bandgap of 1.42 eV, and the $Al_{0.7}Ga_{0.7}As$ region is n-doped with a bandgap energy of 1.82 eV. It is common practice to use upper-case prefixes for heterojunction materials (e.g., N-AlGaAs). The heterostructure introduces a discontinuity in the conduction and valence bands that aids in carrier confinement. The bandgap discontinuity makes it possible to get efficient carrier injection over a very short distance compared to conventional p-n junctions. Although not shown in the figure, the refractive indices of the two semiconductors are 3.59 and 3.39, respectively, which acts to confine photons to the GaAs region with higher refractive index.

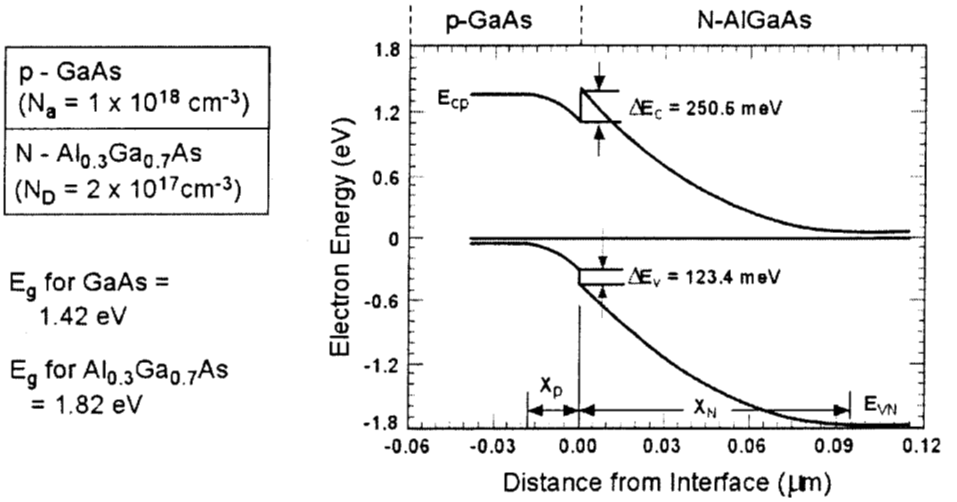



Figure 5-2. Diagram of a GaAs/AlGaAs heterojunction structure.

Materials used for heterojunctions must have lattice constants that are closely matched in order to minimize crystal defects. For example, a lattice spacing mismatch of 1% will introduce a dislocation at approximately every 100 lattice plains (about 500 Å). The requirement for close lattice matching severely limits the material choices for heterojunctions. As discussed earlier, AlGaAs has the unusual property that the lattice constant is nearly the same over a wide range of concentrations, which makes it a good choice for bandgap engineering.

## D. *Strained Lattices in Thin Layers*

Lattice strain is a major problem for conventional heterostructures because it generally introduces defects in the material that interfere with device operation, limiting yield and reliability. However, it is possible to deliberately produce strained lattices over a restricted distance that are stable, and do not result in defects. Strains within such structures change the properties of the crystal in a manner that can be advantageous for optoelectronic devices [New6]. For example, the effective mass of holes and electrons within III-V materials is affected by lattice strain; the effective mass of the holes is considerably reduced in strained lattices, increasing recombination efficiency compared to unstrained lattices of the same material. Auger recombination is also lower in strained lattices. These properties are deliberately exploited to fabricate laser diodes with improved operation at elevated temperature and lower threshold current [New7]. Examples will be discussed in a later section.

E. Quantum Wells

Very compact layered structures can be used to fabricate dissimilar regions that effectively form quantum wells for confined carriers. Quantum effects become significant for layers with thicknesses below 200 A. Figure 5-3 shows the band diagram and energy state densities for a two-level quantum structure. The structure is formed by an AlGaAs/GaAs heterojunction. In this case the width of the structure is less than 100 A, causing the energy diagram in that direction to be restricted to discrete, quantized energy states. The hole states are split into two regions corresponding to "light" and 'heavy" holes.

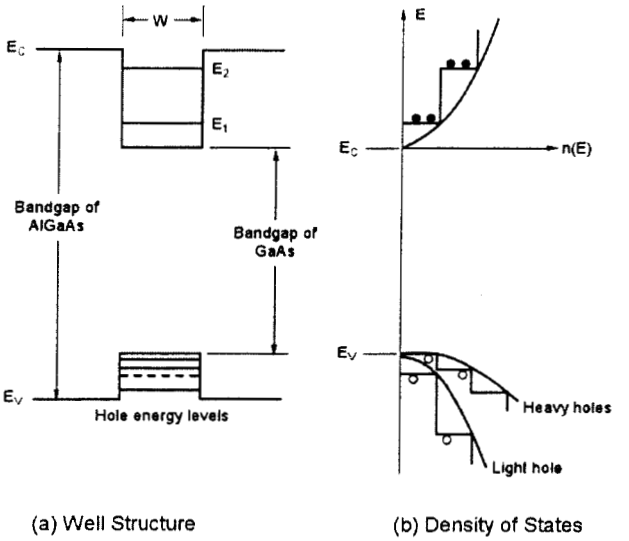



Figure 5-3. Energy band diagram and energy state density of a simple quantum structure.

The quantum properties of the structure modify the absorption spectrum and the effective gain of regenerative optical structures that are used in lasers, increasing the stability of the laser and decreasing threshold current, although it is much more difficult to fabricate lasers with such thin layers. The details are quite complicated, and beyond the scope of this course. Readers are referred to references New 8 and New 9 for details. Many new lasers are fabricated with quantum-well structures, and is important to have a basic understanding of how they are fabricated.

## *6 - Light Emitting Diodes*

### *A. Basic Features*

Light-emitting diodes can be fabricated with direct bandgap semiconductors such as GaAs. The key property is that carrier recombination within the junction formed by the semiconductor regions must have a high probability of producing a photon, along with a long enough carrier lifetime to allow the light to travel to the interface where it can be emitted. Even in diodes made with direct semiconductors most carrier recombination occurs at surface regions under very low injection conditions, and consequently there is very little light output at low current levels. However, when the current through the junction increases to a sufficiently high level the surface recombination saturates and the dominant recombination process begins to occur within the junction. This produces large numbers of photons in a direct bandgap material.

The current-voltage and light power-voltage dependencies of a typical LED are shown in Figure 6-1 (in this figure, the light output was measured with a photodiode, and is measured in relative units). Note the change in slope that occurs in the forward voltage characteristics at the onset of light emission. Ideally the I-V characteristics can be described as the superposition of a non-radiative recombination term (with ideality factor = 2) and a diffusion term corresponding to the region where radiation recombination within the junction dominates (with ideality factor = 1):

$$I = K_{rec} \exp(qV_F/2kT) + K_{diff} \exp(qV_F/kT) \quad (X)$$

where $K_{rec}$ and $K_{diff}$ are the factors for the recombination and diffusion terms, q is electronic charge, $V_F$ is the forward voltage, k is Boltzmann's constant and T is absolute temperature.

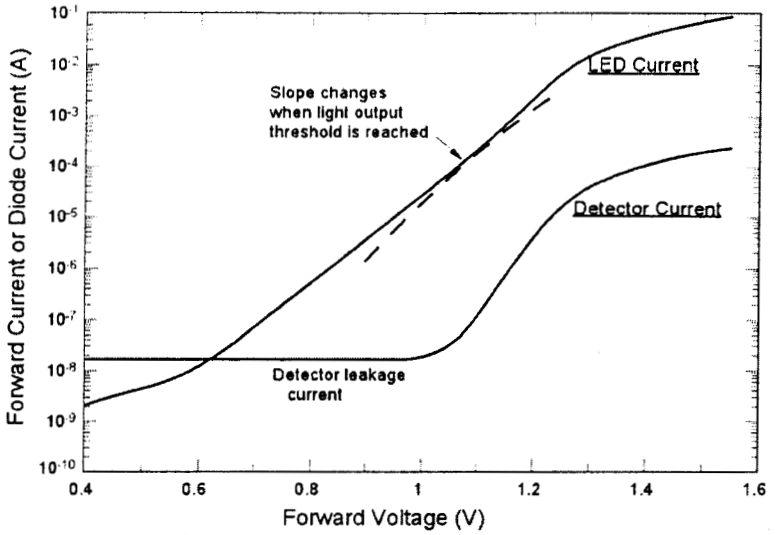



Figure 6-1. I-V and optical power of an LED.

Some types of LEDs show behavior that deviates from the ideal relationship of the previous figure. An example is shown in Figure 6-2 where the slope at low injection changes only gradually as the injection level increases. This may be due to defects within the LED structure that are sensitive to injected current, and also contribute to nonradiative recombination just as the surface recombination component. As noted later in this section, proton damage causes a large increase in the recombination rate at low-to-moderate currents, effectively shunting much of the current into non-radiative processes and decreasing the forward voltage. This causes the slope at low and moderate injection to change in irradiated devices. Thus, even though LEDs are rarely used under low injection conditions, measuring their properties under low injection is an important characterization tool to evaluate non-radiative recombination after radiation damage has occurred. There are also unit-to-unit differences in the slope of the current-voltage characteristics prior to irradiation that may indicate the presence of internal defects in the crystalline structure of the device.

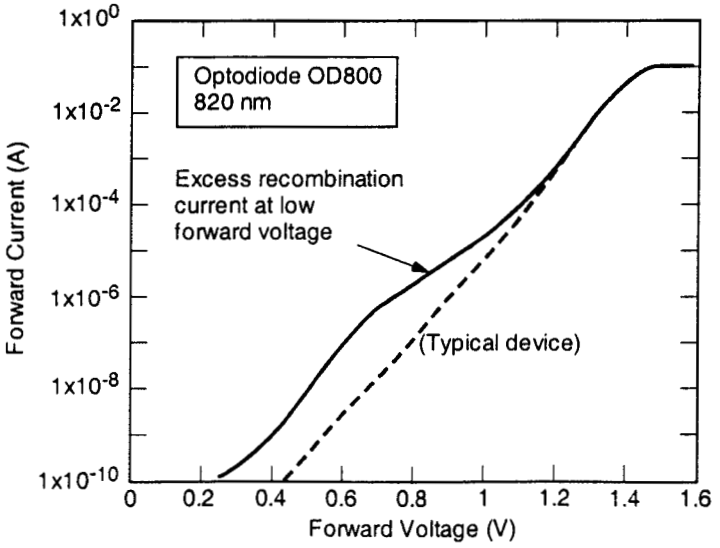



Figure 6-2. Abnormal behavior of I-V characteristics of an LED.

The light output of an LED is affected by temperature, falling approximately 1% for each 1 °C increase in temperature. This is a strong temperature effect, and it must be taken into account when performing radiation tests as well as in determining the degradation and margin in system applications. Many LEDs operate under conditions where the device temperature is 30-40 °C above the nominal temperature of heat sinks or base plates, which causes the light output to be significantly lower than the light output at nominal room temperature conditions.

LEDs also have a "wearout" mechanism that causes gradual degradation in the light output over time (there is no analogous mechanism for conventional silicon electronics). The amount of degradation depends on the LED technology and the amount of current that flows through the LED. Figure 6-3 shows typical degradation curves for wearout for a high-reliability LED that is rated for 100 mA maximum DC operating current. Note the pronounced difference in degradation for devices that are not subjected to initial burn-in conditions. The curves in Figure 6-3 are for typical devices. Wearout degradation can vary for different devices from the same batch or wafer lot. Most LEDs in space applications are operated at currents well below the maximum rated current in order to reduce the amount of degradation from "wearout."

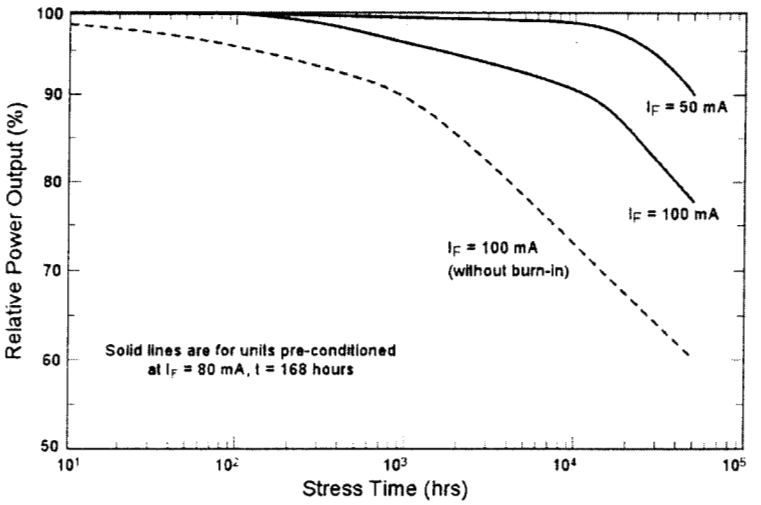



Figure 6-3. Decrease of light output of an LED with extended operation at maximum current.

## *B. Properties of Various LED Technologies*

Many different materials and physical structures can be used to fabricate LEDs, and many advances in LED technology have been made during the last thirty years. In spite of this progress, many LEDs in the near infrared region are still fabricated with a very old process that relies on amphoteric doping, an inexpensive process that produces LEDs with very high output efficiency that are closely matched to the peak responsivity of silicon detectors.

Amphoteric doping relies on a dopant (typical Si for GaAs and AlGaAs) that can act either as an n- or p-type impurity, depending on the growth temperature. With amphoteric doping, it is possible to create a p-n junction in a layer that is initially doped with only a single impurity by gradually altering the temperature during the growth of the epitaxial layer. The resulting structure, shown in the diagram of Figure 6-4, has a graded doping level, and is a compensated semiconductor (the net doping level depends on the difference between the background doping level and the impurities that are altered by the high-temperature growth process). The optical efficiency of amphoteric devices is very high for a number of reasons, including reduced free carrier absorption because of the compensation, and the existence of a complex about 0.1 eV below the valence band that effectively eliminates band-to-band absorption processes that would normally increase the amount of non-radiative recombination, decreasing efficiency [Kres1]. Amphoterically doped devices require relatively long minority carrier lifetimes to operate efficiently because the graded junction extends over a region that is 50-100 $\mu$m wide.

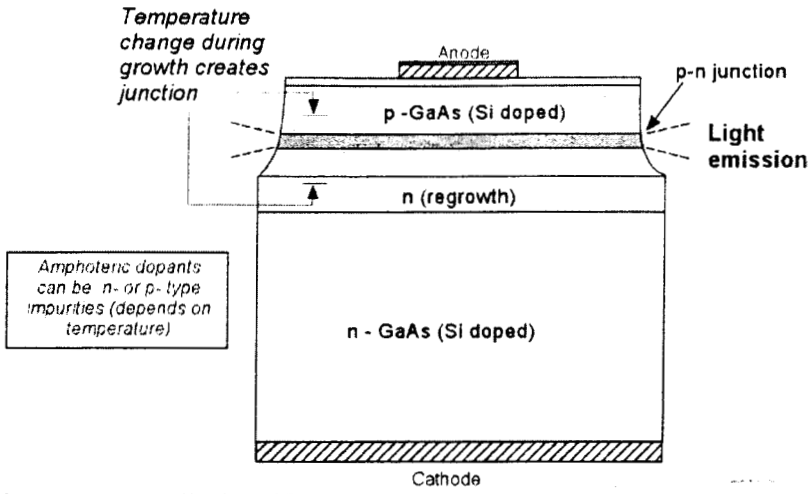



Figure 6-4. Diagram of an amphoterically doped LED.

It is also possible to fabricate LEDs with conventional diffusion processes where different types of impurities are diffused into the starting material to create layers with different doping levels. Most diffused LEDs are made in the visible region, and are generally of less interest for optoelectronic applications in space compared to those in the near infrared region. However, one manufacturer used diffused LEDs with a wavelength of 700 nm in their line of optocouplers, so there are cases where diffused LED performance is important.

More advanced LEDs are manufactured with a multi-layer structure, sandwiching the active region between two layers of a different semiconductor type that confines the carriers to a narrow active region. A diagram of a double-heterostructure is shown in Figure 6-5. Many more fabrication steps are required to fabricate such devices, and the different materials used in the heterojunctions must have nearly identical lattice constants in order to avoid defects. The active region of a double-heterojunction LED is generally very thin, on the order of 1-4 μm. The thin layer allows pulsed or high frequency modulation at higher effective bandwidth than amphoterically doped LEDs, and also makes performance less dependent on lifetime.

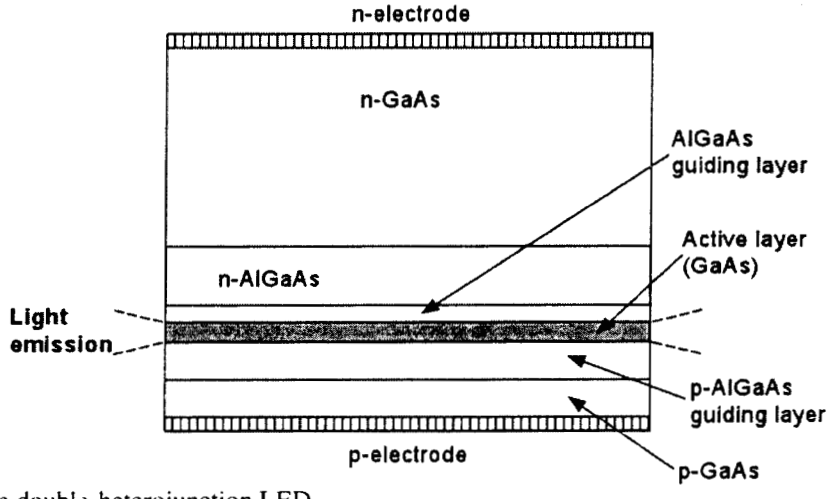



Figure 6-5. Diagram of a double-heterojunction LED.

## C. *Radiation Degradation*

Amphoterically Doped LEDs

The effect of proton degradation on a typical amphoterically doped LED is shown in Figure 6-6. This type of LED is extremely sensitive to displacement damage. Note that the light output has been reduced to about

30% of initial value at a fluence of 3 x $10^{10}$ p/cm$^2$, which is comparable to the total fluence from a single intense solar flare. The damage is higher at low currents, and the current dependence has to be taken into account when radiation testing is done in order to make sure that the actual use conditions are included in the measurement set. Operating the device at high current reduces radiation degradation, it increases the amount of degradation due to aging and consequently most space applications of LEDs restrict the operating current to less than 1/3 of the maximum operating current. Older data on amphoterically doped LEDs also indicated that they were degraded at low levels [Old1, Old2, Old3, Old4], but the measurements were over a more restricted range of currents and operating conditions.

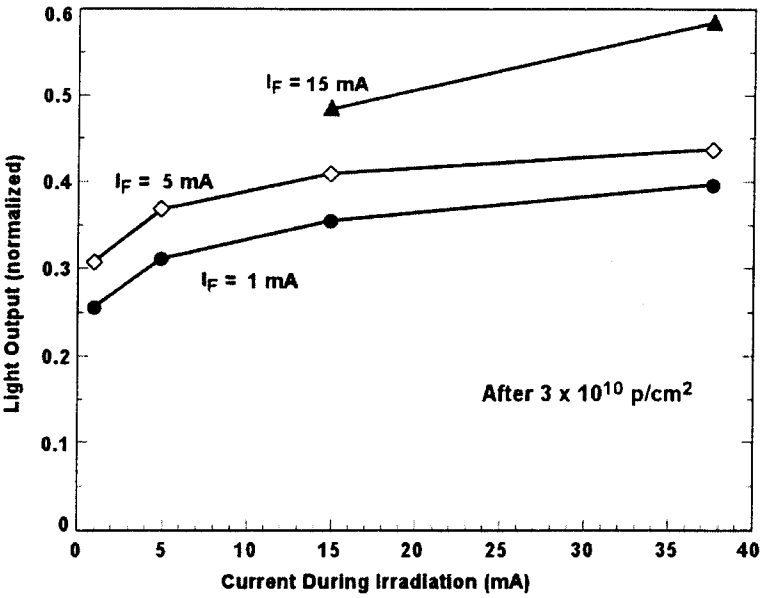



Figure 6-6. Proton degradation of an amphoterically doped LED.

Another way to evaluate LEDs is to measure light output as a function of forward voltage. Figure 6-7 shows a plot of the light output along with a plot of forward current in the diode vs. forward voltage. Note that although the light output is far lower after irradiation, the operating current at which light first begins to be produced is unaffected by the radiation damage. This is in sharp contrast to laser diodes (discussed in the next section) where the threshold current is strongly affected by radiation damage.

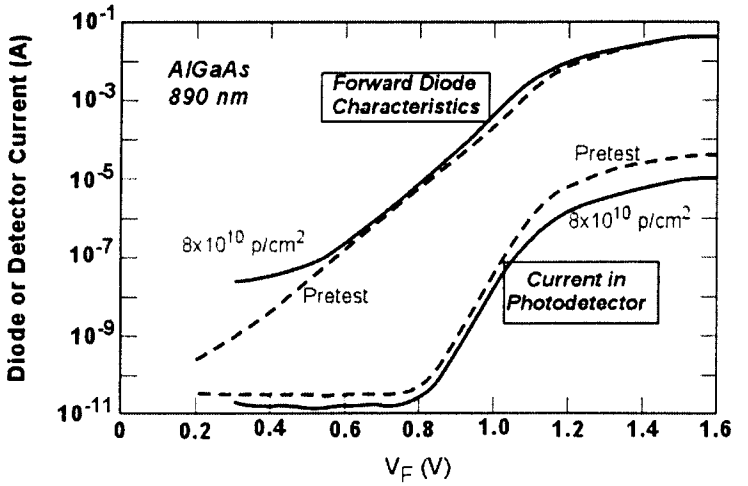



Figure 6-7. Dependence of forward voltage and light output on injection level for an amphoterically doped LED before and after irradiation with protons.

Damage in these types of LEDs is very stable, even over time periods of several months as long as the device remains unbiased. However, part of the damage can be readily annealed by applying a forward current through the device after irradiation [Ann1, Ann2, Ann3]. Figure 6-8 shows the recovery of three different types of amphoterically doped LEDs when a moderate current is applied after they are degraded by radiation. About 20% of the damage recovers if one waits for long time periods, although there are differences in how long one must apply the current to get recovery in LEDs from different processes and manufacturers. The degree of recovery depends on the total charge that flows through the device after irradiation, and is the same even in cases where the device remains unused for periods of many months after irradiation as for devices where current is passed through the device shortly after the irradiation has been completed.

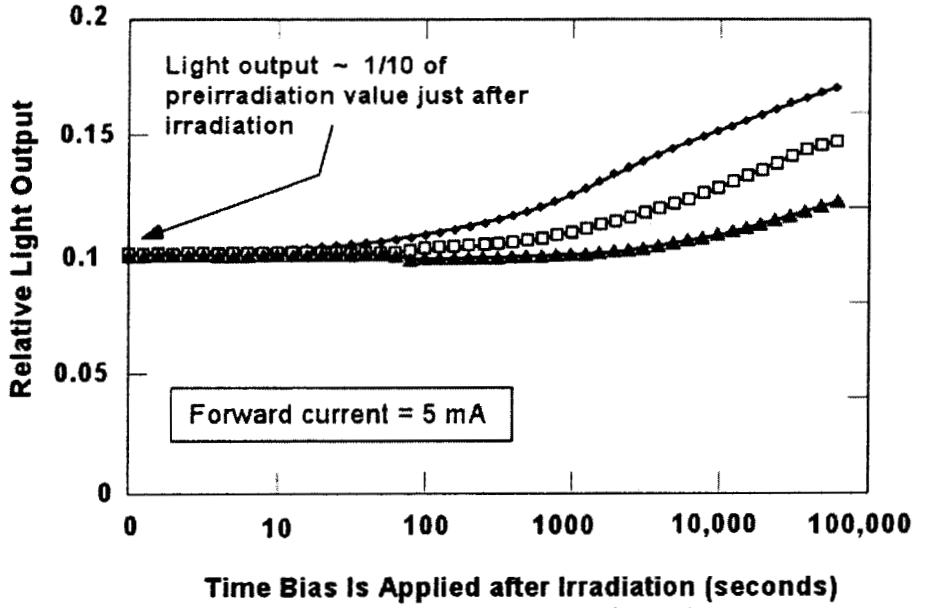



Figure 6-8. Recovery of damage in amphoterically doped LEDs under forward bias.

Diffused LEDs

Although only a limited number of LEDs made with conventional diffused processes have been included in radiation effects studies, they appear to be less susceptible to displacement damage than amphoterically doped LEDs. Most diffused LEDs have shorter wavelengths, within the visible spectrum. Although it is possible to produce conventional diffused LEDs in the 850-930 nm range that is near the peak in responsivity for silicon detectors, amphoterically doped LEDs are much more efficient, and dominate that range of wavelengths.

Figure 6-9 shows the degradation of GaAsP devices with a wavelength of 700 nm when they are irradiated with 50-MeV protons. Although these devices are far less affected by radiation than amphoterically doped LEDs, the initial light output is also lower compared to that of amphoterically doped LEDs in the near infrared region.

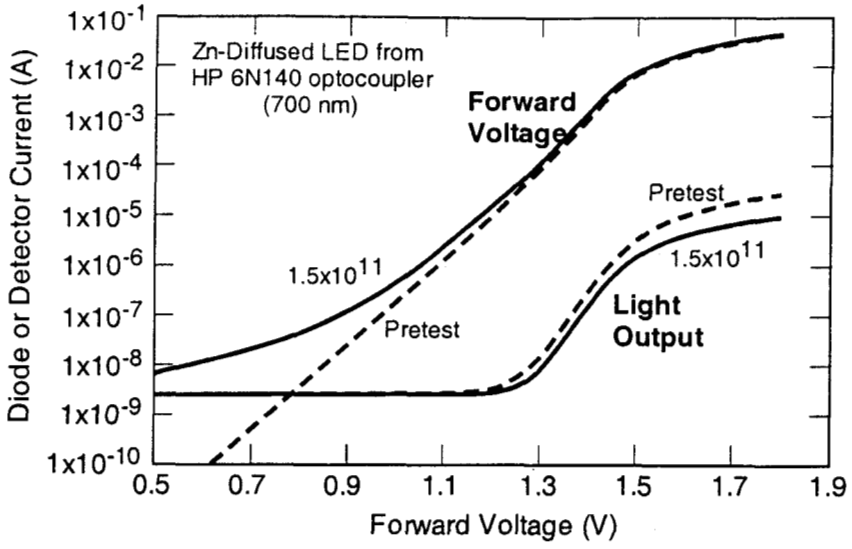



Figure 6-9. Proton degradation of GaAsP LEDs manufactured with a conventional diffused process.

Double-Heterojunction LEDs

Double-heterojunction LEDs are, on average, much more resistant to displacement damage than amphoterically doped devices. Figure 6-10 shows representative results for double-heterojunction LEDs with a wavelength of 820 nm that are intended for high-reliability applications. Unlike amphoterically doped devices, double-heterojunction LEDs do not exhibit injection-enhanced annealing [John1]. Although the increased radiation resistance would appear to make double-heterojunction devices a better choice for space applications, amphoterically doped devices have considerably more initial light output than DH LEDs in the 820-900 nm region, and this tradeoff has to be taken into account when devices are selected and characterized. Device uniformity is another potential issue for double-heterojunction LEDs. In some cases a small number -- 5 to 10% of the population -- exhibit large decreases in light output when they are operated at moderate currents, as shown by the lower set of curves in Figure 6-10. The abnormal devices exhibit a large increase in non-radiative recombination at moderate injection levels, which results in much lower forward voltage after irradiation. This behavior may be related to defects in the material used to fabricate the LED; similar effects have been observed in reliability studies of double-heterojunction LEDs that are subjected to operating stress [New 10].

A comparison of degradation of several different types of LEDs is shown in Figure 6-11. These curves represent mean devices from test lots of approximately 30 parts of each device type. They do not take unit-to-unit variability into account, which is typically about a factor of two for amphoterically doped devices, and somewhat greater (and far less predictable) for double-heterojunction LEDs. The results show that amphoterically doped LEDs degrade by significant amounts at relatively low radiation levels, which can cause severe problems in space applications unless additional design margin is included to take the degradation into account. Visible (diffused) and more advanced double-heterojunction devices can operate at radiation levels that are about two orders of magnitude higher than that of the amphoterically doped devices.

Damage in light-emitting diodes depends on fluence in a nonlinear manner. The relative change in light output is actually greater at higher fluences compared to incremental changes at low fluences. This causes samples with more extreme radiation damage to fall well below the mean values of other devices from the same lot at high fluences. Degradation in amphoterically doped LEDs appears to be dominated by changes in minority carrier lifetime. The levels at which double-heterojunction LEDs degrade is sufficiently high for carrier removal affects to be important, which makes it more difficult to analytically determine how the device degrades. The complex structure of double-heterojunction LEDs, with thin layers of AlGaAs and GaAs material, add further complexity to this problem.

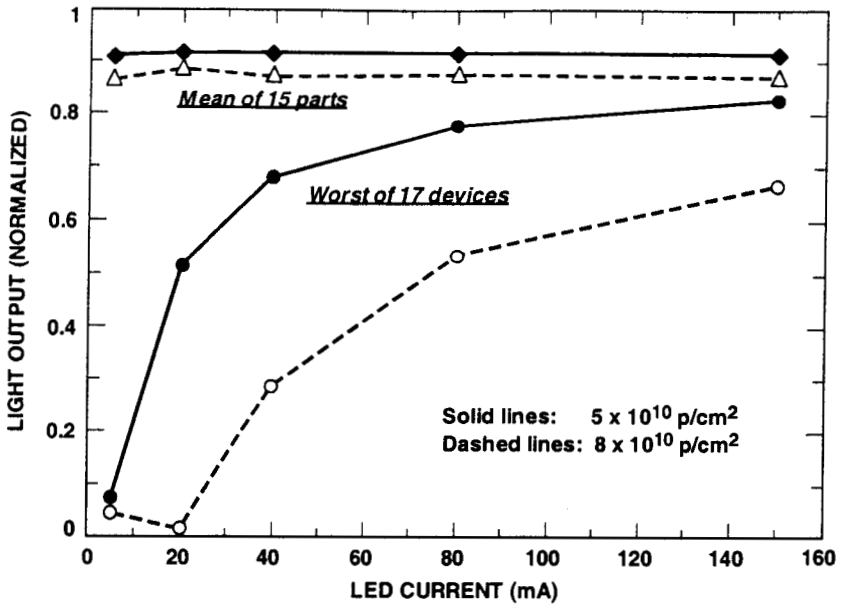



Figure 6-10. Proton degradation of a double-heterojunction LED.

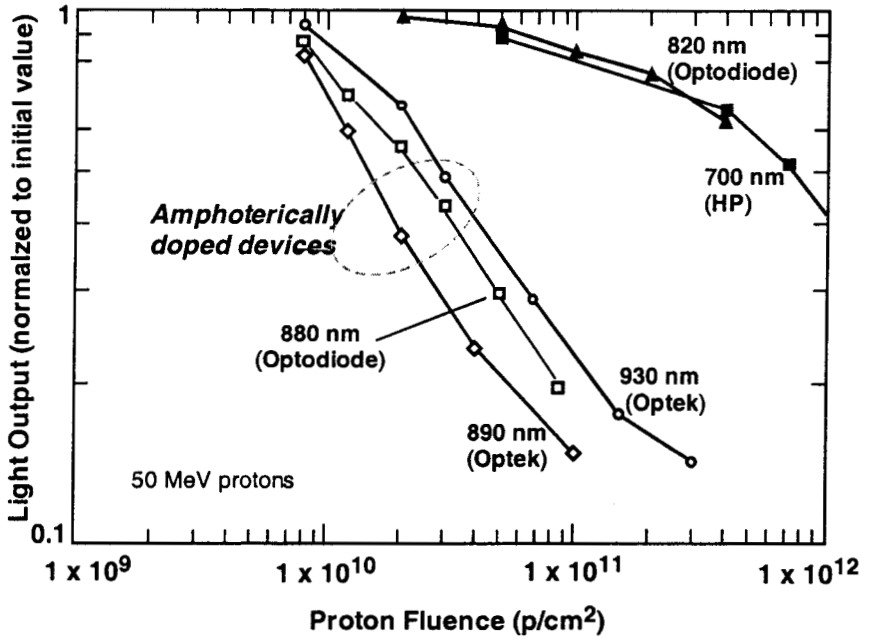



Figure 6-11. Comparison of degradation of several different types of LEDs from various manufacturers.

## *C. Radiation Testing of Light Emitting Diodes*

Proton testing is more difficult and costly than tests with passive sources (such as gamma rays), and considerably more planning is needed to do this type of testing. In addition to facility costs, test boards are activated by the radiation, making it necessary to avoid prolonged exposure to boards and equipment when tests are done between radiation steps.

One approach that can be used for radiation testing is to mount devices in a pattern on a test board that places them within the uniform region of an accelerator beam (typcially a diameter of 5-10 cm). This allows

several devices to be irradiated simultaneously, reducing the testing cost. At predetermined intervals the irradiation is stopped, and the devices are removed from the accelerator area and placed in a light-tight transition fixture that couples light from each LED to a corresponding photodiode. The assembly, shown in Figure 6-11 is designed to provide uniform physical spacing between the LED and the photodiodes used to measure the light output. It may be necessary to control the temperature of the LEDs during measurement to eliminate interference from temperature effects that affect the LED light output. Peripheral electronics are connected to the LED and phototransistor assemblies through cable arrays, using special care to allow low level measurements.

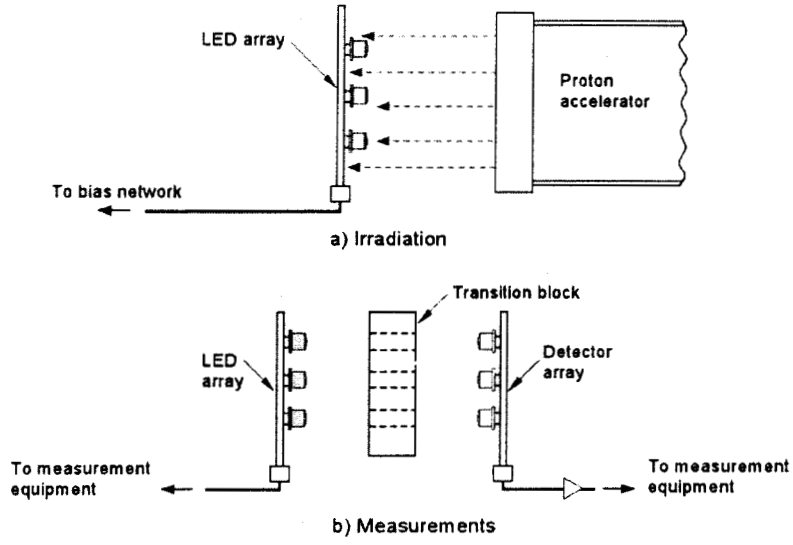



Figure 6-11. Diagram of a test assembly used for irradiation and testing of an array of LEDs.

The light output of LEDs can vary substantially with the position and angle subtended by the detector. This makes it vitally important to have test fixtures that place the LED and detector in fixed, reproducible positions so that the measurements are consistent at different radiation levels (the assembly has to be disassembled for each irradiation because the photodetectors would be damaged by the radiation). Figure 6-12 shows the dependence of light output on angle for two types of LEDs. Note the large "dip" in light output for the LED that is packaged with an internal lens.

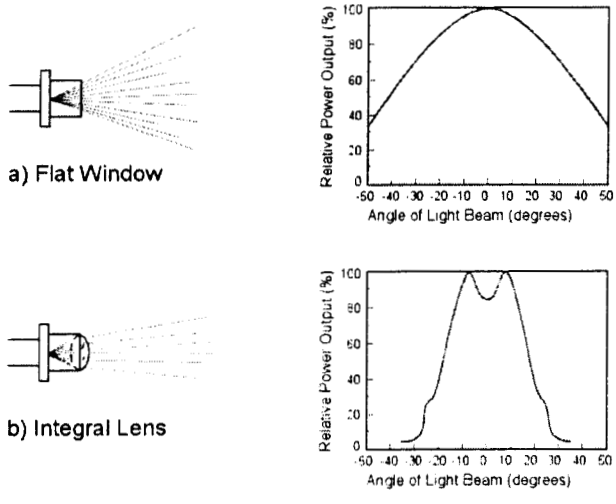



Figure 6-12. Angular dependence of the light output of LEDs in two different package configurations. The LED with the internal lens has a very strong variation in light output with angular position, increasing the difficulty of making reproducible measurements.

The diagram in Figure 6-12 shows another potential issue for LED testing. In some cases the window or lens used as part of the LED assembly can be affected by total dose, which introduces color centers that darken the material. Consequently it is important to evaluate the lens or window, at least indirectly, to make sure that ionization damage in the window or lens is not contributing to the degradation. Alternatively, devices that are tested without an internal lens may underestimate the damage of equivalent devices with an internal lens that is sensitive to total dose darkening.

## *7 - Laser Diodes*

### *A. Basic Features*

Semiconductor lasers depend on the process of stimulated emission. Operation of the laser requires that internal quantum states within the semiconductor are "pumped" by an external or internal source so that their population is well above the thermal equilibrium level (inverted condition). Stimulated emission occurs when a photon -- produced as a result normal recombination processes -- travels through the pumped laser cavity. The initial photon will cause an additional photon to be produced through recombination while producing a new photon that is an exact duplicate of the photon that initiated the recombination process. A property of stimulated emission is that *the duplicate photon has the same energy, direction, and polarization as the initial photon* [New 11]. For an optical cavity with an inverted population, the probability of stimulated emission becomes very high. The result is a "flood" of photons with nearly constant energy and direction that are triggered by a small number of initial photons (produced by normal recombination processes) within the cavity. Steady-state laser operation requires that the rate of production of excited levels equals the rate that they are depleted by stimulated emission, leading to a minimum threshold current condition for operation. Laser cavities have an effective optical gain that depends on the material properties and the design of the laser cavity [New 12]. The threshold current (or current density) required to establish operation of the cavity as a laser is inversely proportional to the optical gain.

Most semiconductor lasers are internally pumped, relying on current that flows through a p-n junction within the laser structure to pump the internal laser cavity with photons. As discussed above, these photons are generated by recombination within the junction. Figure 7-1 shows a simplified structure of a double-heterojunction laser. In this example the active layer is GaAs, surrounded by p- and n-doped layers of AlGaAs that form a p-n junction in a transverse direction to the laser cavity. The AlGaAs layers have a lower refractive index, confining the photons that are emitted by stimulated emission to the GaAs region at the center. The AlGaAs layers form heterojuctions that injects carriers very efficiently into the GaAs region. The laser is similar in construction to the double-heterojunction LED, but has cleaved facets to partially reflect the light in a precise transverse direction along the axis. The laser facets must be parallel and smooth, which is an additional constraint during manufacturing. It must also be designed to withstand higher current densities than typical LEDs.

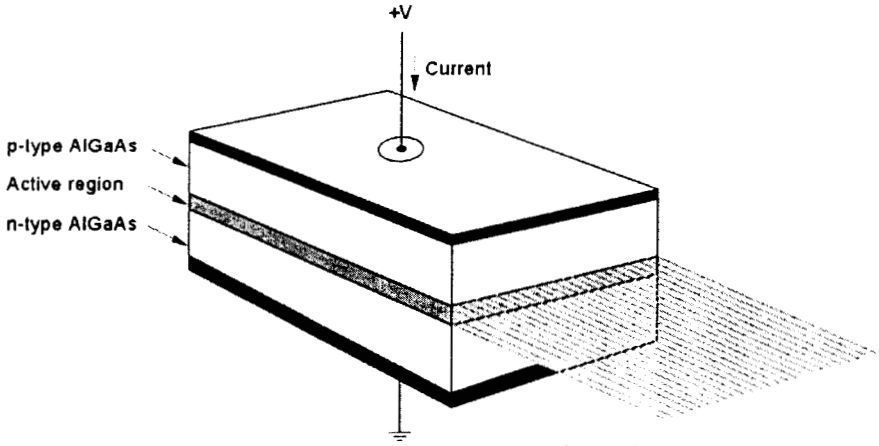



Figure 7-1. Simplified diagram of a double-heterojunction semiconductor laser.

The dependence of light intensity on injected current is shown in Figure 7-2 for an older laser technology. At low injection, the slope of the forward current vs. $V_F$ characteristic is two. In this region most of the recombination occurs at surface regions or at non-radiative defect sites within the junction, and very little light is produced. As the current increases, the slope of the current-voltage curve decreases to one, and the structure begins to emit light because most of the recombination is now due to radiative recombination within the junction. In this region the laser operates very much like a light-emitting diode because the optical gain of the cavity is too low for laser operation. The spectral width is relatively broad, typically 6-8% of the peak wavelength

As the current increases, the slope begins to change once again, and it increases with a very steep slope when the injection level is high enough to cause the structure to lase (at approximately 100 mA in this example). Once this region is reached, there is an abrupt increase in light power along with a sharp decrease in spectral width to a very small fraction --0.2 to 1% -- of the peak wavelength. At still higher current the slope decreases due to thermal effects and internal resistance. Under pulsed conditions the slope at high current is considerably steeper (heating is reduced), and it is possible to operate the device at much higher currents compared to steady-state conditions.

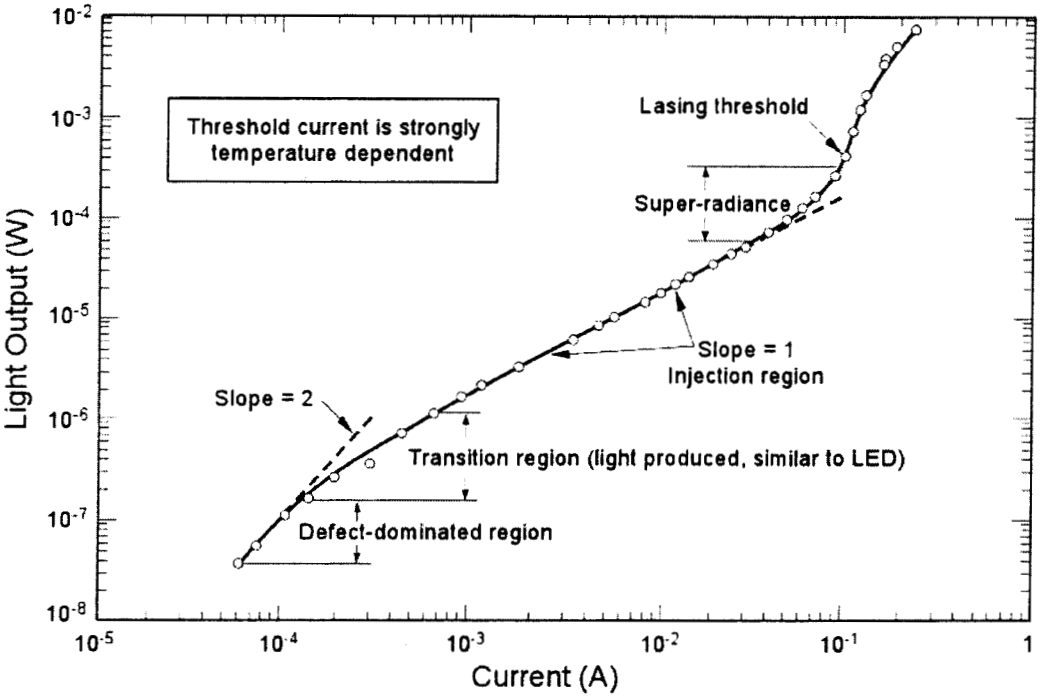



Figure 7-2. Light output vs. injection level for a laser diode and LED.

Threshold current is one of the most important parameters, and the threshold current of newer laser structures is considerably lower than the older example shown in Figure 7-2. Threshold current is highly sensitive to temperature, as shown in Figure 7-3, and for AlGaAs it also depends on the composition. Because of the extreme sensitivity to temperature it is usually necessary to precisely control the laser temperature, or to provide external feedback circuitry to control the current through the laser at fixed value above the threshold current. This can be done by monitoring the light output, and using the sampled light as the input to the feedback control network. Some lasers are fabricated with internal photodetectors that can be used for that purpose. Lasers operate at much higher operating current densities than LEDs, which can adversely affect their reliability. This has been a barrier to their application in space systems in the past, but laser technology has improved to the point where they can be used in undersea applications and long-haul communications networks with comparable reliability requirements [New 13]. Nevertheless, it is more difficult to use lasers than LEDs in most systems, partly because of the need to carefully control the current through the device and to limit the actual operating temperature range.

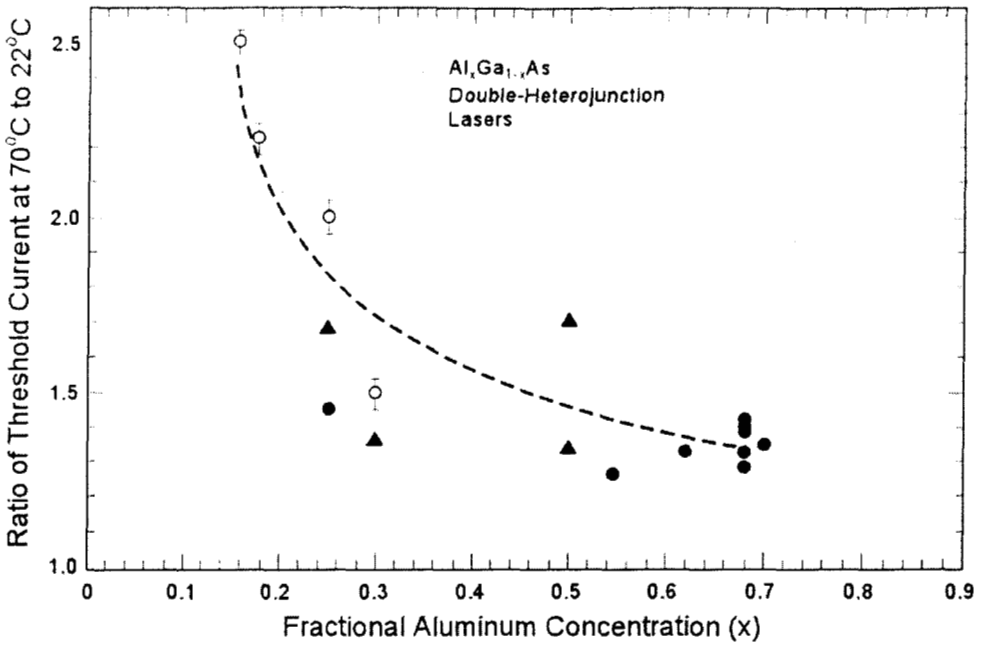



Figure 7-3. Dependence of threshold current on temperature.

Lasers operate at very high internal power levels, and they generally have more severe wearout and reliability problems compared to LEDs. Laser failures can be grouped into three different categories:

(1) gradual increase in threshold current with extended operation

(2) gradual degradation of light output due to the development of dark-line defects

(3) catastrophic degradation or "steps" with lower light output, attributed to facet degradation because of the high power levels associated with laser operation.

Lasers fabricated with AlGaAs/GaAs (with wavelengths between 940 and 990 nm) exhibit all three failure modes. However, lasers that are fabricated with the InGaAs/InP system (1300 -1500 nm) appear to be less affected by dark line defect degradation than lasers fabricated with AlGaAs/GaAs. Changes in material technology have lowered the defect level, and those changes along with decreases in threshold current density have improved laser reliability considerably compared to reliability of earlier lasers.

## *B. More Advanced Laser Structures*

Numerous advances have been made in laser technology during the last 20 years [Las1-Las5], and these changes have allowed special structures to be made with new materials, covering a much wider range of wavelengths compared to older laser diode structures. Threshold current has also been reduced by several orders of magnitude for newer types of lasers. The ability to grow extremely thin layers of material by molecular-beam epitaxy provides many additional degrees of freedom in designing specific laser structures. However, this is very confusing when one examines the relatively limited data on radiation damage that is available for lasers, because the laser structures are often quite different.

The first lasers were homojunction structures which required extremely high current densities, ≈ 100,000 A/cm$^2$. They were only capable of operation for a few hours. Heterojunction structures were developed in the early 1970's which provided a much more efficient way to inject carriers and decrease threshold current density. Table 7-1 shows how the threshold current has decreased during forty years of development. Note how much the threshold current has been reduced.

Table 7-1. Threshold Current Trends in Laser Development

| Technology | Year | Threshold Current ($A/cm^2$) | Active Region Thickness (Å) |
|---|---|---|---|
| Homojunction | 1965 | ≈ 100,000 | 25,000 |
| Single Heterojunction | 1968 | 12,000 | 4,000 |
| Double Heterojunction | 1970 | 1,600 | 1,800 |
| Quantum Well | 1980 | 500 | 120 |
| Strained Quantum Well | 1990 | 65 | 60 |

Although it is not possible to cover laser technology in a great deal of depth in this course, we will briefly discuss strained layers, which are used in some types of lasers. Properly designed strained layers can be used to design lasers with improved operation. Some of the advantages of strained layers include reduced effective mass for holes (the high effective mass of holes in unstrained semiconductors limits efficiency), reduced Auger recombination, and the ability to operate the laser structure at higher temperatures [New Ref]. Figure 7-4 shows a calculation of the optical gain of an InGaAs laser structure (designed for a wavelength of 1.55 μm) for an unstrained lattice and one that is deliberately designed to induce a strain of 1.5% along one axis of the crystal [New 14]. The optical gain is nearly four times higher in the strained layer, with a current density that is also much lower. This illustrates how strained lattices can be used to design lasers with lower threshold current and improved operation.

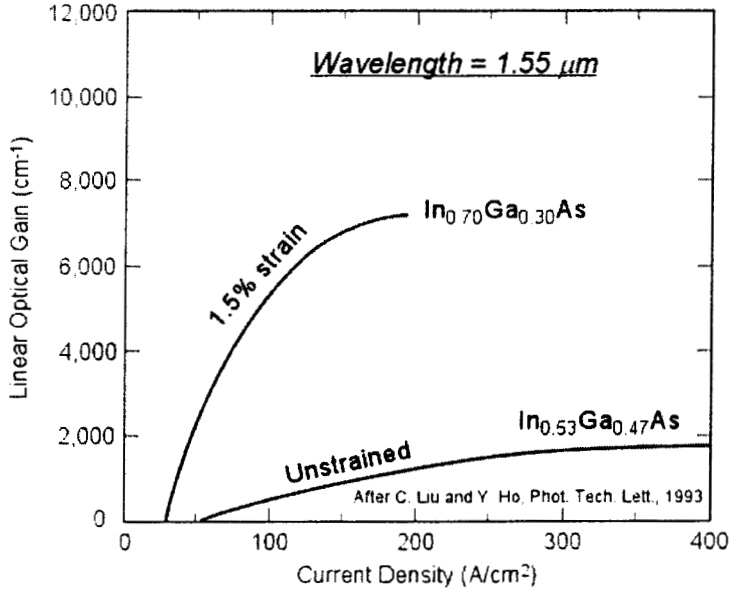



Figure 7-4. Optical gain of strained and unstrained InGaAs laser cavities.

Strained layers can be combined with quantum wells to design advanced lasers. Single and multiple quantum-well lasers are available that have far better performance than older laser technologies. However, these structures require very sophisticated processing steps along with the ability to grow extremely thin layers of material that are closely matched to the wavelength of the laser. The thickness and structure of the layers used in fabrication have to come very close to meeting the conditions for the expected laser wavelength in order for the structure to work properly, and this may cause larger unit-to-unit differences to occur in the properties of the lasers and their sensitivity to radiation damage compared to older structures used in LEDs and lasers.

## C. *Radiation Degradation*

Only a limited amount of data is available on radiation degradation of lasers, partly because earlier laser technologies did not have adequate reliability for applications in military or space systems. The operating life of early lasers was on the order of hours. The reliability of newer lasers is far better, but it is still more difficult to apply lasers in space because of the need to control operating current and temperature. Lasers are now being considered for space use, particularly in high-speed data links.

Threshold current is one of the most critical parameters of semiconductor lasers. Displacement damage causes the threshold current of a typical laser to increase, as shown in Figure 7-5 for an older AlGaAs laser [Chow1]. In this example the device (tested in 1989) was irradiated with neutrons; it is shown because there is very little data available with protons for older laser structures. Note the very high threshold current. Moderate levels of radiation damage shift the threshold current to higher values, but cause only a very slight change in the slope of the light output characteristics (sometimes referred to as slope efficiency). At higher levels the nature of the damage changes, and the slope is severely degraded along with the threshold current. The increase in output power at high currents after the highest radiation level is due to annealing from the high power dissipation within the laser. Although this data is for neutrons, NIEL calculations show that 1-MeV neutrons have nearly the same effective displacement damage effect as 200 MeV protons, allowing comparison with more recent laser structures with proton degradation.

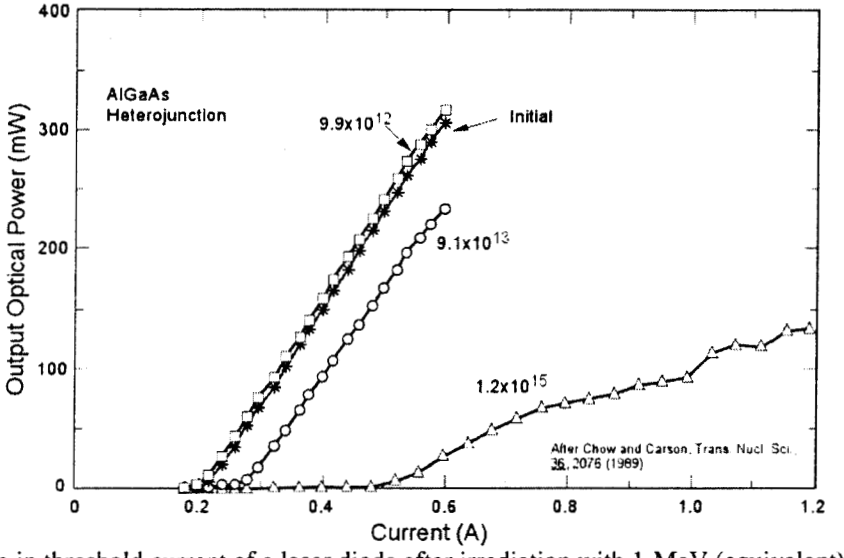



Figure 7-5. Increase in threshold current of a laser diode after irradiation with 1-MeV (equivalent) neutrons.

An example of degradation of a more advanced laser structure is shown in Figure 7-6 [Evan1]. This was an advanced strained quantum-well laser that was designed for operation over a wide temperature range. The original data was taken with 5.5 MeV protons because they were conveniently available to the research group that developed the laser. The data in the figure have been altered by applying a NIEL factor of 6.4 so that the results are equivalent to 50 MeV protons. This was done to allow a more intuitive comparison of the degradation of this device with other data in the course. The main effect of the radiation damage is to shift the threshold current; the magnitude of the change is almost exactly proportional to fluence. Note that the slope is essentially unchanged, even after the highest radiation level.

More recent data on a commercial multi-quantum well laser (operating at 780 nm) is shown in Figure 7-9 [Zhao1]. The results are quite similar to those obtained for the strained quantum-well laser in the previous figure, although the threshold current of the 780 nm device is slightly lower. Threshold current shifts by about the same relative amount after irradiation, and there is little change in the slope efficiency. In subsequent work they reported significant variability in the radiation degradation of different units from the same test lot [Zhao2], illustrating that unit-to-unit variability can be an important issue for advanced lasers.

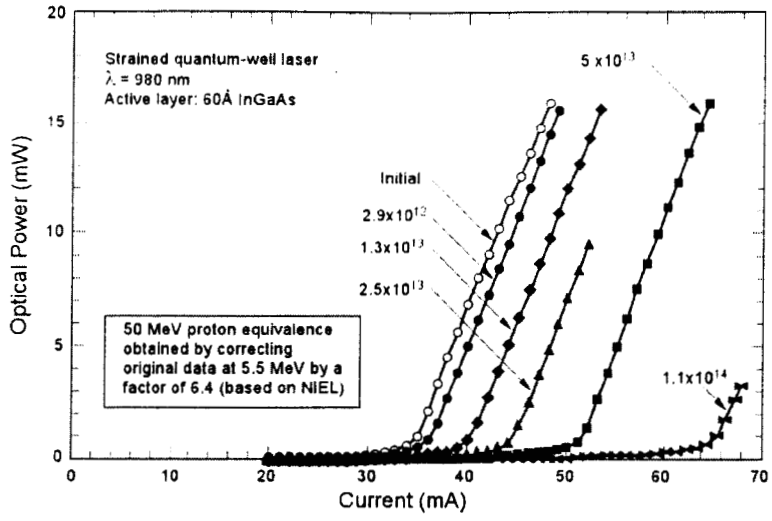



Figure 7-6. Degradation of a strained quantum-well laser after irradiation with protons (data reported in equivalent 50 MeV protons, applying a factor of 6.4 to allow for the increased effective damage of the 5.5 MeV protons used in the original work).

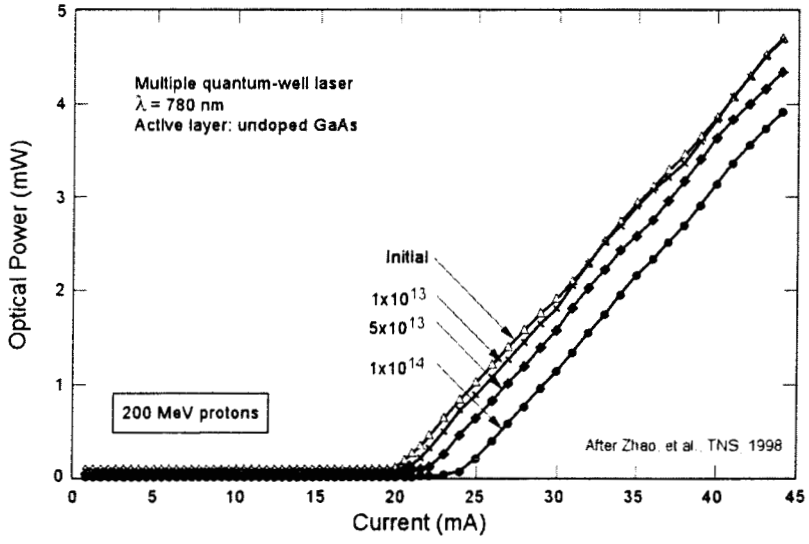



Figure 7-7. Degradation of a multi quantum-well laser after irradiation with protons.

Zhao et al. also found that the lasers that they studied were sensitive to injection-enhanced annealing. In some cases the threshold current that was initially degraded by radiation recovered to a *lower* value than that observed prior to irradiation. They attributed this to change in the refractive index of internal ridge waveguides used in the laser structure, although other effects may also contribute. Their results illustrate yet another complication in evaluating advanced types of lasers.

## *D. Vertical Cavity Semiconductor Lasers (VCSELs)*

A new approach for laser design has been developed during the last ten years that provides many advantages compared to conventional laser diodes. A diagram of a vertical-cavity semiconductor laser (VCSEL) is shown in Figure 7-8 [Vcsl1, Vcsl2]. VCSELs have very low threshold currents compared to conventional laser structures because of the small dimensions of the laser cavity. VCSELs can be made with very low threshold current, a major advantage compared to conventional laser diodes. Their small size also reduces the

dependence of threshold current on temperature, which not only reduces the power required for the device, but also eliminates the need for external temperature control or feedback circuitry that is usually required for more conventional lasers. VCSELs are usually operated over a relatively narrow range of currents because internal heating of the small laser cavity reduces output efficiency at currents that are more than about a factor of three above the threshold current.

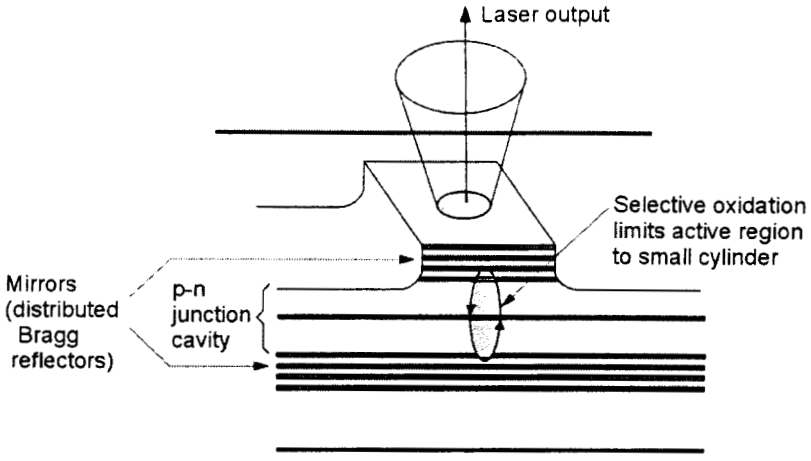



Figure 7-8. Diagram of a vertical-cavity semiconductor laser.

Proton damage in VCSELs causes the threshold current to increase, just as for conventional lasers. Figure 7-9 shows results from VCSEL devices that were fabricated at Sandia National Laboratory [Vcsl3, Vcsl4]. Note that the slope efficiency of VCSELs is more affected by moderate levels of radiation than for conventional lasers. At high currents the small VCSEL cavity undergoes substantial heating, increasing the amount of non-radiative current in the device. That is the main reason for the large drop in light output at higher operating conditions after irradiaton. The limited radiation data on VCSELs indicates that they also have somewhat larger unit-to-unit variability in radiation degradation compared to optical emitters that do not incorporate such complex construction.

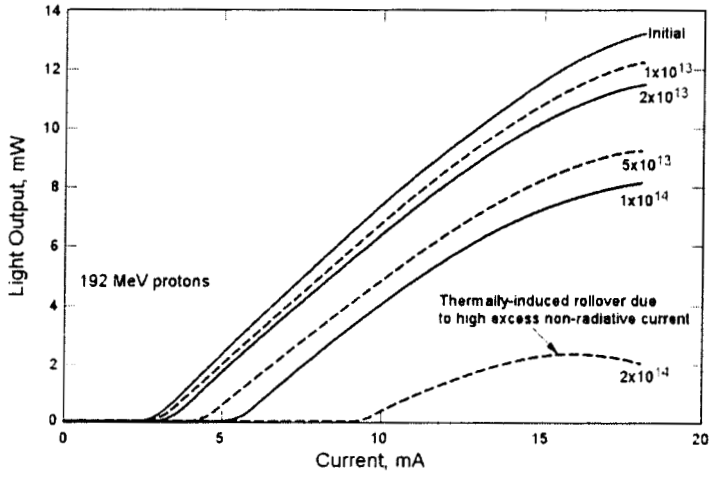



Figure 7-9. Effect of protons on the operating conditions of a VCSEL.

## *E. Radiation Testing Considerations for Semiconductor Lasers*

The basic approach used for characterizing lasers is similar to that of LEDs, with even more concern about physical alignment between the laser and detector because lasers typically have a much narrower beam angle. Special care needs to be taken to control the device temperature with high precision when measurements are

made between irradiations. This is required for lasers because the threshold current is so strongly affected by temperature that the temperature dependence will interfere with measurements to determine the effects of radiation on threshold current except for very large changes. Special thermoelectric modules are available that provide a convenient way to control temperature for devices that are suitably mounted. Although controlling case temperature is effective when moderate currents are used for characterization, internal heating can be important at high currents and it may be necessary to restrict the operating duty cycle to keep from heating the device to the point where it affects measurements.

The most important parameters for lasers are the threshold current and slope efficiency. Lasers are often intended for fiber-optic applications, and it may be necessary to extend the measurements to much higher current values using pulsed current sources (with careful attention given to controlling the power and duty cycle) in order to encompass the operating characteristics in the application within the measurements done to characterize radiation degradation. The data in Figures 7-7 and 7-8 shows that the optical power can be more degraded at high operating currents, and it is not possible to extrapolate data taken at lower currents to the very high currents that are sometimes used for pulsed operation.

More thorough characterization testing can be done, including measurement of wavelength and spectral width. This requires an optical spectrometer, and such measurements are considerably more time consuming than indirect measurements of the optical power with a photodiode or other basic detector. At moderate radiation levels most lasers that have been tested in the past do not show significant changes in wavelength or spectral width, but this may not be valid for new types of lasers. The underlying mechanisms for degradation of advanced laser technologies are not well understood. Carrier removal and microscopic cluster damage, similar to that reported for silicon CCDs [Srou2] may play a role in damage in new types of lasers, and more thorough characterization is generally recommended.

## *8 - Optical Detectors*

### *A. Basic Considerations*

Several different mechanisms can be used to detect photons, but the treatment here will be limited to two mechanisms that are useful with semicocondutor detectors: photoconductivity and the photovoltaic effect.

Photoconductivity

Light absorbed within a semiconductor reduces the electrical conductivity because excess carriers are generated (recall that conductivity, $\sigma$, is related to the carrier density by the equation $\sigma = n q \mu$ where n is the carrier density, q is electronic charge, and $\mu$ is the mobility). For intrinsic semiconductors, absorption depends on wavelength as discussed in Section 3, up to the bandgap edge (the photon energy must exceed the bandgap energy). Table 8-1 lists the properties of some semiconductors that are commonly used as detectors.

Table 8-1. Properties of Some Semiconductors Used as Detectors

| Material | Temperature (°K) | Wavelength Limit (nm) |
|---|---|---|
| CdS | 295 | 0.52 |
| GaP | 295 | 0.56 |
| GaAs | 295 | 0.92 |
| Si | 295 | 1.1 |
| InGaAs | 295 | 1.6 |
| Ge | 295 | 1.8 |
| PbS | 295 | 2.9 |
| InAs | 195 | 3.2 |
| PbSe | 195 | 5.4 |
| InSb | 77 | 12 |
| HgCdTe | 77 | 12 |

Normally semiconductor detectors rely on direct excitation of carriers from the valence band to the conduction band, and can only be used at wavelengths below the "bandgap edge." However, if the semiconductor is doped with a very high number of impurities then there is a very high concentration of impurity levels *within* the bandgap. and it is possible to have transitions from the impurity states to the valence band that correspond to much lower energies compared to direct transitions. This process is referred to as extrnisic photoconductivity. For example, indium can be used as a dopant to extend absorption in silicon to about 5 $\mu$m [IR1]. This type of detector needs to be operated at temperatures below 77 K in order to improve the signal-to-noise level

Measurements (or applications) of photoconductivity require a suitable electrical circuit, and thus photoconductivity actually consists of the combination of the excess carrier generation process along with the way that the excess carriers are transported. Photoconductivity is the mechanism that causes photocurrents in reverse-biased photodiodes, discussed in the next subsection.

Photovoltaic Effect

Excess carriers generated by absorbed light produce photocurrents that will develop a voltage across an initially unbiased p-n junction. One application of the photovoltaic effect is in solar cells, which are designed to provide relatively large currents to external sources. It is also possible to use the photovoltaic effect for low-level photodetectors. That mode of operation provides some improvement in signal-to-noise ratio compared to photoconductive detection, but is generally slower than photoconductive processes in semiconductors because the low electric field that is present causes most of the charge to be collected by diffusion.

## *B. Elementary Detectors Based on p-n Junctions*

Responsivity

The responsivity of detectors is determined by the dependence of the absorption coefficient on wavelength, discussed in Section 4-C, along with the effective depth of charge collection in the p-n junction structure. The lightly doped material must extend well beyond the maximum absorption depth in order to collect light near the bandgap edge. For indirect materials, such as silicon, this requires a collection depth of 100 $\mu$m or more. However, photodiodes made with direct bandgap materials do not require this extended depth because the absorption coefficient changes very little until the wavelength reaches the wavelength corresponding to the bandgap edge. Figure 8-1 shows the responsivity of a typical deep p-n silicon detector along with that of an AlGaAs detector (the wavelength limit can be tailored by selecting different concentrations of aluminum).

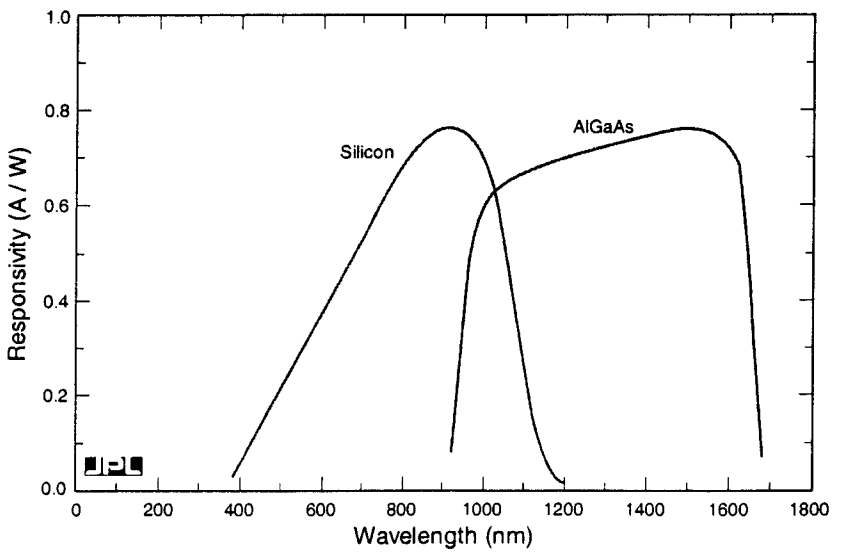
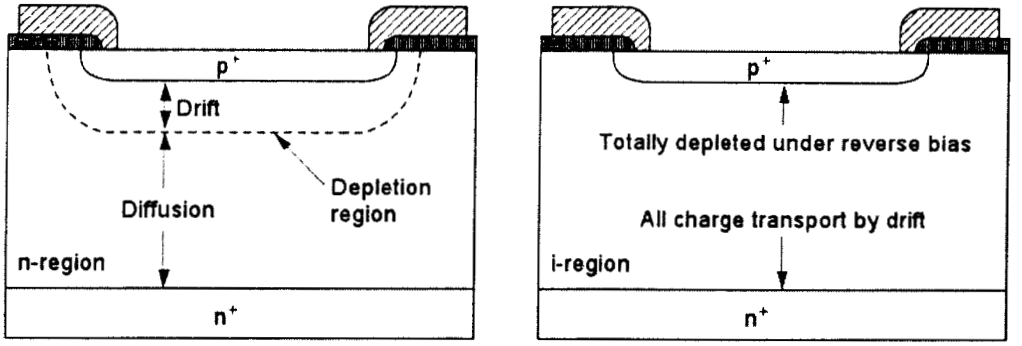



Figure 8-1. Responsivity of silicon and InGaAs detectors.

p-n Junction

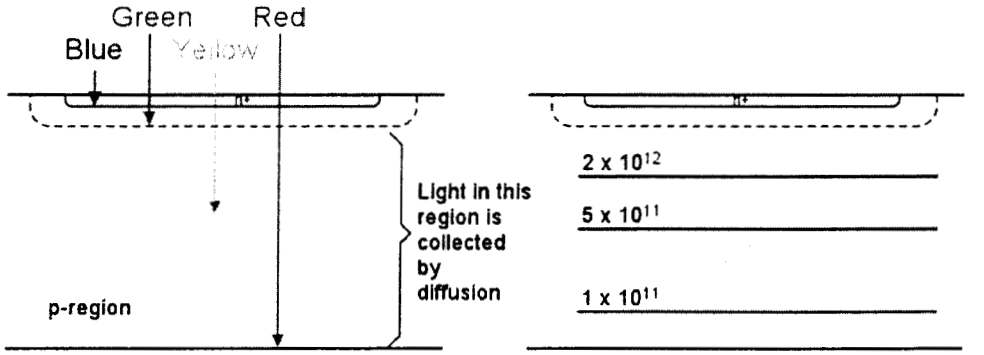
Photodiodes are one of the most commonly used photodetectors. When used in the reverse-biased mode, excess carriers generated within the depletion region are collected very rapidly because of the strong electric field that is present. Carriers generated outside the depletion region are collected by diffusion. This is shown schematically in Figure 8-2(a).




(a) Conventional Photodiode (b) p-i-n Photodiode

Figure 8-2. Excess carrier generation and collection in a biased p-n junction.

The depth over which light is collected in a photodiode depends on the absorption coefficient (see Section 3.x). This is depicted for a silicon photodetector (for which $\alpha$ has a somewhat gradual dependence on wavelength) in Figure 8-3. Light at very short wavelengths is collected at the surface, and much of the photocurrent produced in that region is lost because of recombination in the highly doped contact region. Light at longer wavelengths is absorbed deep within the material. For example, at 850 nm the absorpion depth is about 40 $\mu$m, and the depth of the detector must extend beyond that value in order to efficiently collect light at longer wavelength. Charge collection at longer wavelengths are more affected by radiation damage because the diffusion length is reduced, as shown in the figure.




(a) Absorption by Various Wavelengths (b) Diffusion Length After Irradiation with 50-MeV Protons: $L = \sqrt{D\tau}$

Figure 8-3. Diagram showing change in absorption depth for light of different wavelengths in a silicon detector.

Phototransistor

It is possible to use photocurrent in the collector-base region of a transistor as a p-n detector and then use the gain of the transistor to amplify photocurrent. There are several advantages in such structures, including reduction of stray capacitance and improved response time. A physical diagram of a phototransistor is shown in Figure 8-4. The structure consists of an extended base region surrounded by a narrow emitter ring (the triangular region of the emitter is a contact region provided to attach a bonding lead). The extended open base

region functions as an integrated photodetector. Although the physical diagram is for a single transistor, a compound (Darlington) transistor is often used that provides considerably more gain compared to that of a single transistor.

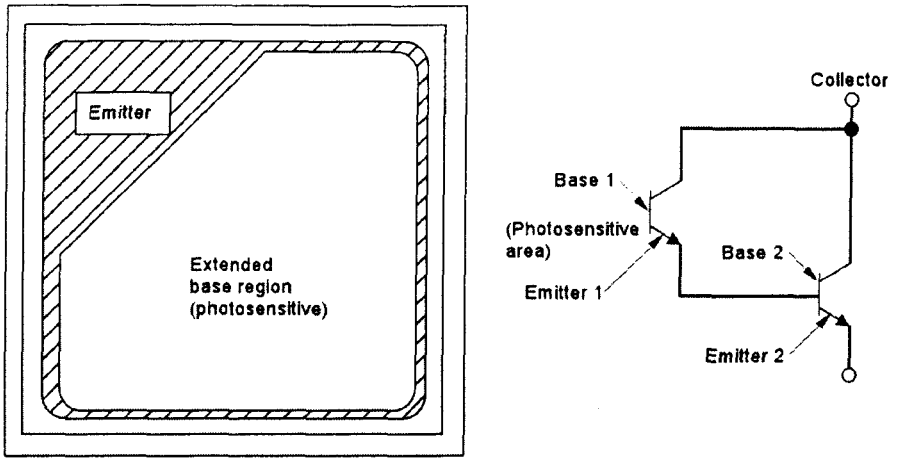



Figure 8-3. Physical structure of a typical phototransistor.

The bandwidth limitations that are inherent in a p-n junction detector because of charge collection by diffusion can be overcome by interposing a wide lightly doped intrinsic region between the p- and n-regions. Such a p-i-n diode structure was shown in Figure 8-1, along with the conventional photodiode structure. When a sufficiently high reverse bias is applied, the entire i-region is depleted of carriers and consequently all the photocurrent generated within the i-region is collected by field-enhanced depletion, not diffusion. This increases the response time and also results in improved signal-to-noise ratio.

It is possible to manufacture p-i-n detectors from direct-bandgap III-V semiconductors as well as from silicon. As discussed in Section 4-C, direct-bandgap detectors can be made with a very shallow collection volume. This allows high-speed detectors to be fabricated that can be fully depleted at lower voltages compared to silicon p-i-n detectors.

Radiation Damage in Conventional and p-i-n Detectors

The photoresponse of silicon p-i-n detectors is much less affected by radiation damage than the photoresponse of conventional p-n photodiodes, as shown in Figure 8-4. This difference is particularly large at

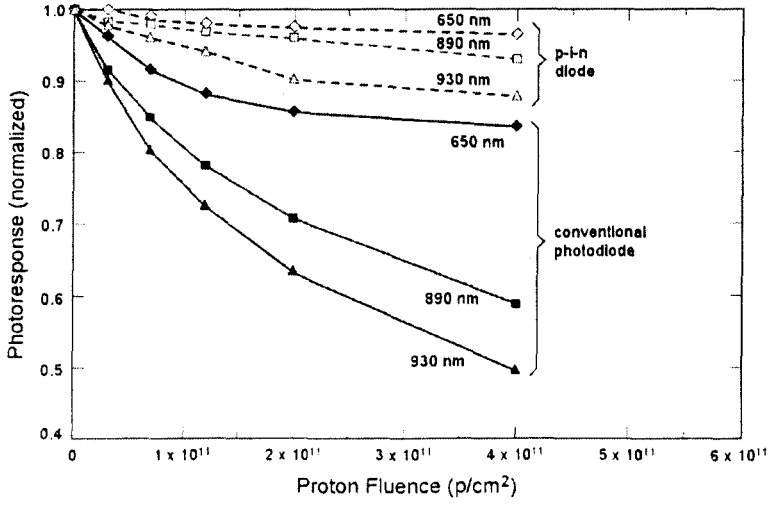



Figure 8-4. Effects of proton radiation damage on the photoresponse of a conventional and p-i-n detector.

longer wavelengths near the maximum respoonsivity of silicon, and is due to the fact that light collection in p-i-n detectors does not depend on minority carrier lifetime. However, carrier removal effects in the lightly doped i-negion cause leakage current to increase, and these changes are noticeable at levels below $10^{10}$ p/cm$^2$. Leakage current changes in conventional photodiodes are much smaller because the doping levels are higher.

## *C. Specialized Detector Technologies*

### Avalanche Photodiode

Avalanche photodiodes are physically very similar to conventional p-n photodiodes. However, they are operated at electric fields that are within the avalanche breakdown of the p-n junction where avalanche multiplication acts to increase the photocurrent. It also affects leakage current. The avalanche factor can be a factor of ten or more, providing higher current and a significant improvement in signal-to-noise compared to conventional photodiodes. However, the bias conditions must be carefully controlled to keep the device in a stable operating mode. A diagram of an avalanche photodiode is shown in Figure 8-5. The solid line shows the I-V characterisitics without light, and the dashed line shows how photocurrent affects the characteristics.

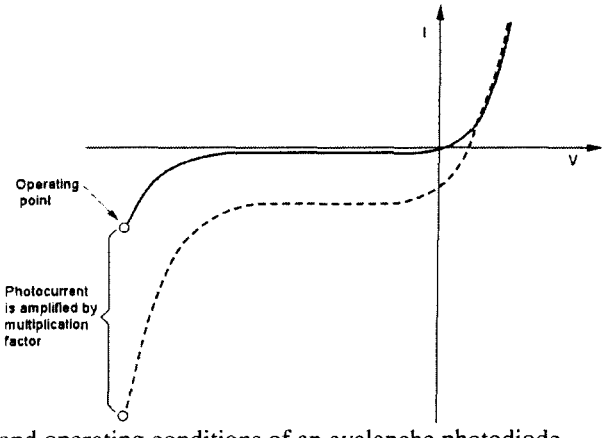



Figure 8-5. Physical diagram and operating conditions of an avalanche photodiode.

### MSM Photodetectors

Another type of detector, the metal-semiconductor-metal (MSM) detector, is also used for high-speed applications, such as fiber optic links. MSM detectors are essentially Schottky diodes. This type of detector was discussed in detail in Part 3 of the 1999 Short Course by P. Marshall [Mars2], and that reference should be consulted for more details.

## *D. Noise and Figures of Merit*

### Basic Noise Concepts

Noise occurs in resistive materials in the form of a fluctuating voltage or current. The mean square noise power $P_N$ of a measurement is given by

$$P_N = k\,T\,\Delta f$$

where k is Boltzmann's constant, T is absolute temperature, and $\Delta f$ is the bandwidth of the measurement. Real noise measurements usually result in measurement of voltage or current with a specific source resistance, R. The ideal voltage noise $v_N$ from an unbiased resistor is then given by

$$v_N = (4\,k\,T\,R\,\Delta f)^{1/2}$$

which has the somewhat unusual units of volts/$\sqrt{\text{Hz}}$. However, this form of unit arises naturally when noise is discussed. Note that for an ideal resistor the noise power density is the same at all frequencies. Noise from a resistor arises from basic thermodynamic considerations, and it establishes the lower limit for noise with a

given bandwidth and resistor value. It is sometimes referred to as Johnson noise. An equivalent form to equation (X) for the noise current that arises from the noise power can be easily derived.

A second form of noise is generation-recombination (G-R) noise. For an intrinsic photoconductor where G-R noise is dominated by only one type of carrier (typically electrons), the short-circuit current term that occurs because of G-R noise from a current $I_B$ is

$$I_N = \frac{2I_B}{N} \left[ \frac{P \tau \Delta f}{1 + \omega^2 \tau^2} \right]^{1/2}$$

where N is the number of free electrons, P is the number of free holes, B is the bandwidth, $\tau$ is the lifetime of free carriers, and $\omega$ is the angular frequency.

A third form of noise, shot noise , occurs because of fluctuations in the flow of carriers within a semiconductor junction. The shot noise current $I_N$ due to a current I flowing through the junction is given by

$$I_N = [(2 q I + 4 q I_o) \Delta f]^{1/2} \qquad (Y)$$

where $I_o$ is the ideal reverse bias junction current in the diode equation.

A fourth noise source also exists that is termed "1/f noise" because it rises rapidly at very low frequencies.

An example of how these noise sources affect noise in an actual low-noise amplifier is shown in Figure 8-6. The importance of the various noise contributions depends on the bandwidth and frequency range that is of interest. The first component, Johnson noise, dominates for high frequencies or for detector applications with wide bandwidth. For more limited bandwidth applications the other noise terms become dominant. At very low frequencies, 1/f noise is usually the dominant noise term. Note that all noise terms are proportional to the square root of the bandwidth, and thus it is necessary to restrict the bandwidth in order to make effective low level measurements in cases where the sensitivity limit approaches the ideal noise "floor". This figure is a useful guide in determining which noise sources are likely to be the dominant problem in low-noise designs.

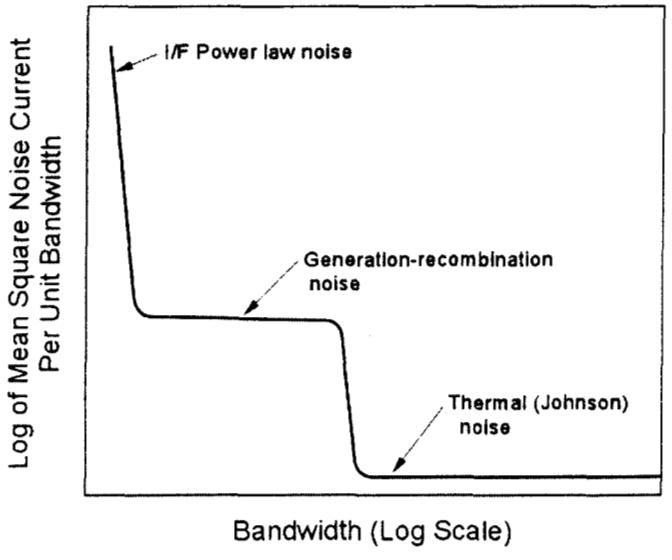



Figure 8-6. Relative contributions of the different noise terms for an actual detector.

Signal-to-Noise Ratio: D*

Noise in infrared detectors is still more complicated because they are sensitive enough to detect blackbody radiation from passive structures that surround them as well as the more usual electrical contributions to noise. Thermal energy from the background (or from materials that are within the field of view) ultimately limits sensitivity. A different figure of merit D* has been developed for detectors that are intended to operate near maximum sensitivity limits, and that figure of merit is usually used for infrared systems instead of the equivalent circuit noise terms that were discussed above. D* is defined as the rms signal-to-noise ratio in a 1-Hz bandwidth per unit rms incident radiant power (per square root of detector area), with units of cm-$Hz^{1/2}$/W. D* can be defined in two different ways, either in response to a monochromatic source, usually referred to as $D^*_{\lambda}$,, or in response to a reference source of blackbody radiation (a wide range of wavelengths). In the latter case, the reference temperature is typically 500 K. The detector field of view is assumed to be hemispherical. The equation below shows how D* is related to normal measurement parameters

$$D^* = \frac{(A_{det}\,\Delta f)^{1/2}\; v_s}{P \qquad v_n}$$

where $A_{det}$ is the detector area, $\Delta f$ is the bandwidth, P is the incident radiant power, $v_s$ is the rms signal amplitude and $v_n$ is the rms noise amplitude.

Figure 8-7 shows D* for various wavelengths and fields of view, assuming a background temperature of 290 K. At short wavelengths, there is very little difference in D* for different fields of view, but there is a very large dependence on viewing angle for wavelength between 5 and 10 $\mu$m. Far better detector performance can be obtained in the infrared region by cooling the detector and any associated apertures or baffling. More details can be obtained from Part 4 of the 1993 Short Course [Pick1], including how various types of infrared detectors are affected by radiation.

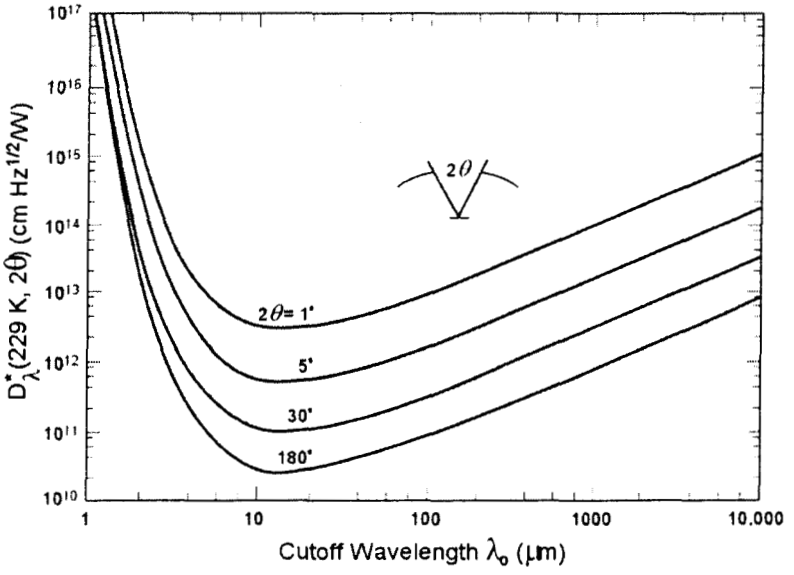



Figure 8-7. D* for various wavelengths and fields of view, assuming a background temperature of 290 K.

Circuit Issues

One circuit technique that is frequently used for photodetectors is the transimpedance amplifier, shown in Figure 8-8. This circuit uses a high-gain, low noise operational amplifier as a current-to-voltage converter. It provides much higher bandwidth than conventional amplifiers that use a resistor in series with a detector to provide an input voltage, partly because to first order there is no voltage change at the amplifier input (virtual

ground). That lowers the effective capacitance at the input. Special transimpedance amplifiers are available that provide bandwidths above 1 GHz for high speed applications.

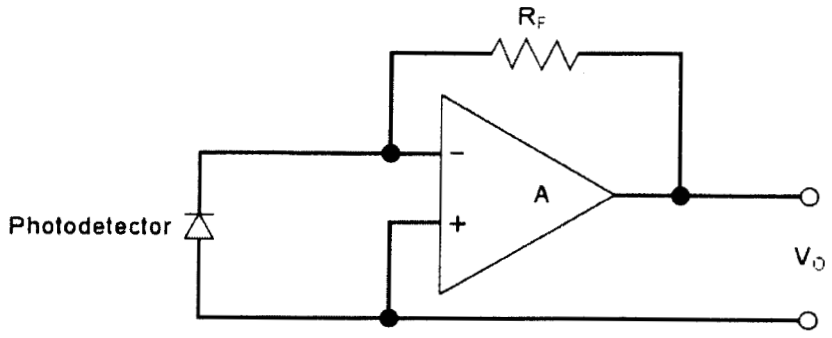



Figure 8-8. Diagram of a transimpedance amplifier. Transimpedance amplifiers are frequently used in detector applications because they provide higher bandwidth and lower noise compared to conventional amplifiers.

In most cases better performance can be obtained by integrating a preamplifier (or complete amplifier) with the photodetector, much like the phototransistor is integrated in Figure 8-3. One of the main advantages with integrated amplifiers is the marked reduction in capacitance between the detector and amplifier input. Noise in high frequency amplifiers increases with the 3/2 power of input capacitance [Pers1].

## *9 - Optocouplers*

### *A. Basic Features*

Optocouplers use a light emitter (typically an LED) to provide an internal optical signal to a photodetector and amplifier (or phototransistor). The optocoupler provides a very high degree of isolation between the electrical signal that drives the LED and the output of the amplifier because there is no direct electrical connection between them (other than very small stray capacitance), only the optical signal. Figure 9-1 shows two different construction techniques for optocouplers. Optocouplers are very simple hybrid devices, consisting of an LED assembly, mounted on a carrier, with a silicon integrated circuit containing a photodiode and transistor (or high-speed amplifier). Some manufacturers produce only the photodiode/amplifier, purchasing the LED from outside sources, while others fabricate -- and control -- the LED as well as the silicon-based part of the optocoupler.

Different physical configurations are used to fabricate optocouplers. Direct coupling, shown in Figure 9-1(a), uses a surface-emitting LED that is inverted and placed directly over the photodiode in the silicon die. This approach is straightforward, providing highly efficient light coupling. A thin layer of optical coupling material (barely detectable in the figure) is usually placed between the LED and phototransistor in order to reduce Fresnel losses.

The indirect method shown in Figure 9-1(b) uses a side-emitting LED. It relies on total internal reflection from a silicone compound that is placed over the LED and detector/amplifier. Optocouplers with indirect coupling are easier to fabricate compared to those with the direct coupling method shown in Figure 8-1(a). However, the amount of light that is transmitted depends on physical properties that are difficult to control -- the roughness of the cleaved edge of the LED and the presence of bubbles in the silicone -- as well as on the electrical properties and optical efficiency of the two materials. Many optocouplers are made with indirect coupling because of the ease of manufacturing them along with improved voltage isolation between the LED assembly and the silicon subassembly compared to optocouplers with direct coupling.

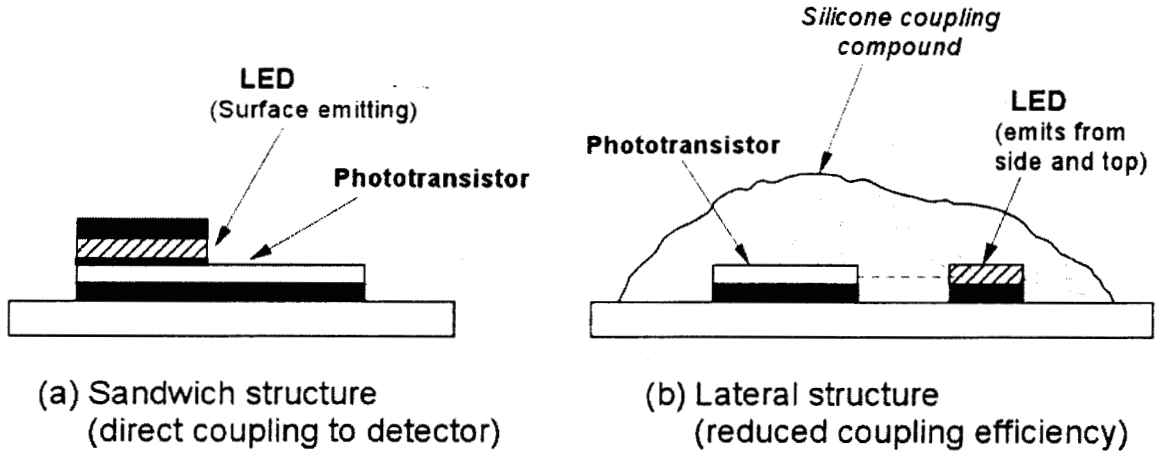



Figure 9-1. Diagram of two basic optocoupler configurations.

One of the most important parameters of an optocoupler is the current-transfer ratio (CTR) which is the ratio of the output current of the amplifier to the forward current in the LED. For a simple optocoupler with a transistor output, the CTR is closely analogous to the gain ($h_{FE}$) of a bipolar transistor. However, the optical process is relatively inefficient, resulting in CTR values between 1 and 10.

Optocouplers can be designed for either digital or analog applications. The negative temperature dependence of the LED causes the CTR to vary over a relatively wide range. For digital optocouplers, the variability can be dealt with quite easily by requiring the LED to be driven well beyond the "active" CTR region (essentially specifying the device only for a saturated condition). Figure 9-2 shows the I-V characteristics of a simple digital optocoupler with a transistor output stage. The characteristics are very similar to that of a discrete transistor, with the LED forward current taking the place of base current to the transistor. The dashed curve for each input current show how the transfer characteristic change when the device is heated to a modest temperature, 36 °C, compared to the lower curve at 25 °C. The curves decrease at higher temperature because the LED output drops with temperature much faster than transistor gain increases with temperature.

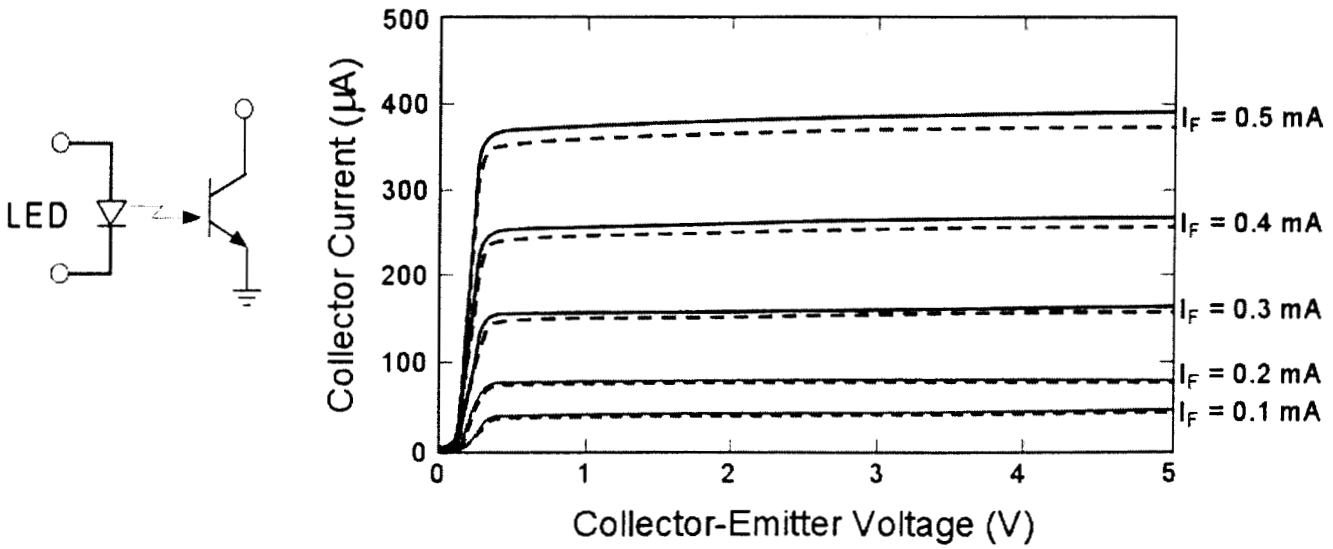



Figure 9-2. I-V characteristics of a digital optocoupler with a simple tranasistor output stage showing effect of heating on characteristics.

Transfer characteristics of a more complex digital optocoupler with an integrated amplifier and digital output stage are shown in Figure 9-3. The "active" region where the device makes the logical transition is typically a factor of three or more lower than the conditions under which logical operation is guaranteed by the device specifications. The logical "low" specification drives the LED current to the point that the output is strongly saturated (note the break in the horizontal axis scale). The active switching point also depends on output loading, as shown by the dashed line, but this dependence is sublinear because of the high gain of the internal amplifier.

Measurements of the digital output in saturation provide no information about where the device actually makes the logical transition, and effectively prevent one from determining how parametric changes in the LED output and amplifier photoresponse are affected by radiation. Digital measurements also "mask" unit-to-unit variability in the active switching region, which can be important. Measurements of the active switching region are recommended as auxiliary measurements for radiation characterization even though this is a special measurement that is not included in the standard device specifications. These measurements are easily made for most digital optocouplers, although some types of amplifiers contain internal comparators with hysterisis that require two sets of measurements, one for high-to-low, and the other for low-to-high transitions.

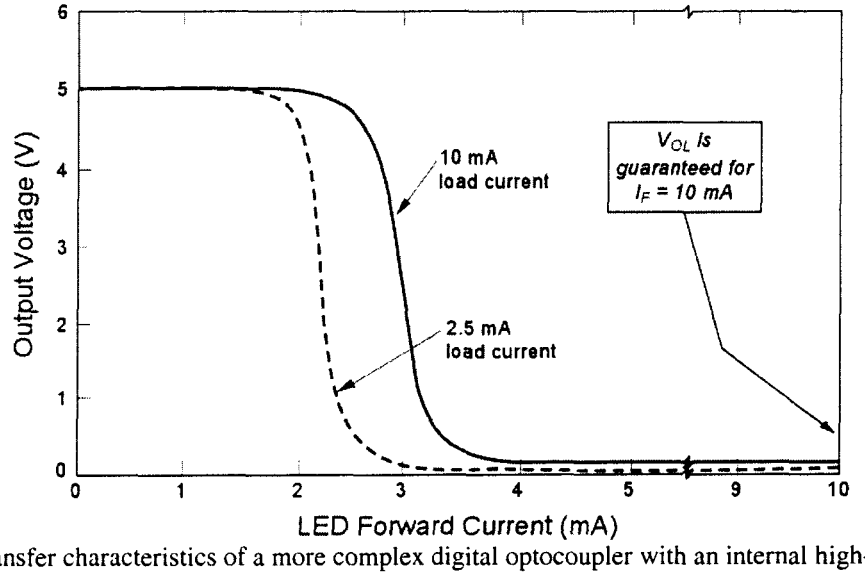



Figure 9-3. Transfer characteristics of a more complex digital optocoupler with an internal high-gain amplifier.

The design and operation of analog optocouplers is quite different from that of digital optocouplers. Analog optocouplers require that the CTR be held within very close limits over a broad temperature range (this is not an issue for digital optocouplers for which temperature effects are "swamped out" by simply overdriving the device into saturation). Analog optocouplers usually have a very basic phototransistor (or photodiode/transistor) output stage. One way to reduce the temperature dependence of such a structure is to operate the transistor in the high-injection region, well above the maximum point in the $h_{FE}$-$I_C$ curve of the phototransistor. In this region, the output current actually increases slightly when the (effective) base current provided by the photocurrent from the LED decreases, compensating for the drop in LED output at higher temperature. This design approach allows the current transfer ratio to be kept within very narrow limits, even over a wide range of temperatures. It also makes analog optocouplers inherently less sensitive to changes in LED output due to radiation until the changes are large enough so that the phototransistor operates at low injection instead of high injection.

## *B. Radiation Degradation*

Digital Optocouplers

The 4N49 is an example of a basic digital optocoupler that has frequently been used in space or other high-reliability applications. This device uses a simple phototransistor as a detector/amplifier, along with an LED with unspecified characteristics (the manufacturer can select any wavelength or LED type that will enable to meet the overall electrical specifications of the optocoupler). The CTR is directly related to the transistor gain and the LED output. The three manufacturers that provide the 4N49 all use amphoterically doped LEDs; two manufacturers do not manufacture the LED, but obtain it from outside sources.

The data in Figure 9-4 show how CTR of the 4N49 is degraded by 50-MeV protons. Note the extreme sensitivity of this device to displacement damage, which is mainly due to degradation of the LED. The protn fluence from a single intense solar flare (~ $10^{10}$ p/cm$^2$) is sufficient to degrade the CTR by about a factor of three when it is operated at low forward current. The recommended forward current of the 4N49 for high-reliability applications is 1 mA because of concerns about LED wearout. That current is far below the maximum operating current, and inadvertently makes the device considerably more sensitive to radiation damage. The damage is lower for conditions where higher current is used for the LED, even when the device is irradiated without bias and with a low duty cycle to minimize annealing. The lower damage is the result of operating the phototransistor at higher injection levels.

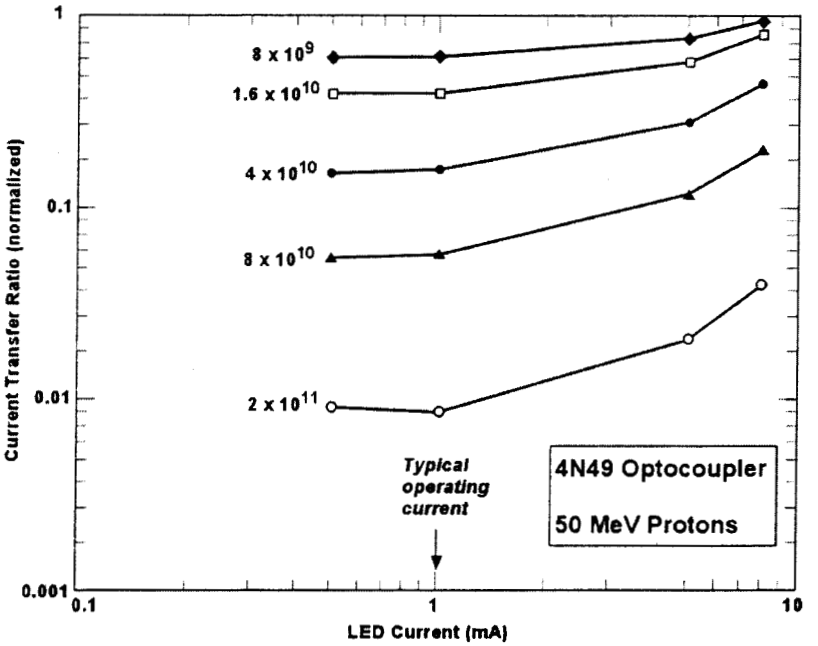



Figure 9-4. Proton degradation of a widely used optocoupler with an amphoterically doped LED.

It is possible to examine the different factors that control CTR in the 4N49 separately by partially disassembling the device and measuring the light output of the LED, transistor gain, and photoresponse separately after each irradiation. Note that an external light source of constant amplitude was used so that radiation degradation of the photoresponse is made at the same light injection level, not the lower injection level that occurs when the internal LED is the light source The results are shown in Figure 9-5 [Rax1]. Note that the CTR degrades more severely at low currents than would be estimated from the product of the photoresponse and LED output. This is because the phototransistor operates at lower and lower current levels when the LED output degrades. The phototransistor operates less efficiently under those conditions, substantially increasing the overall degradation. Phototransistor gain contributed very little to the degradation. The optocoupler CTR degrades by more than two orders of magnitude at the highest radiation levels used in these tests.

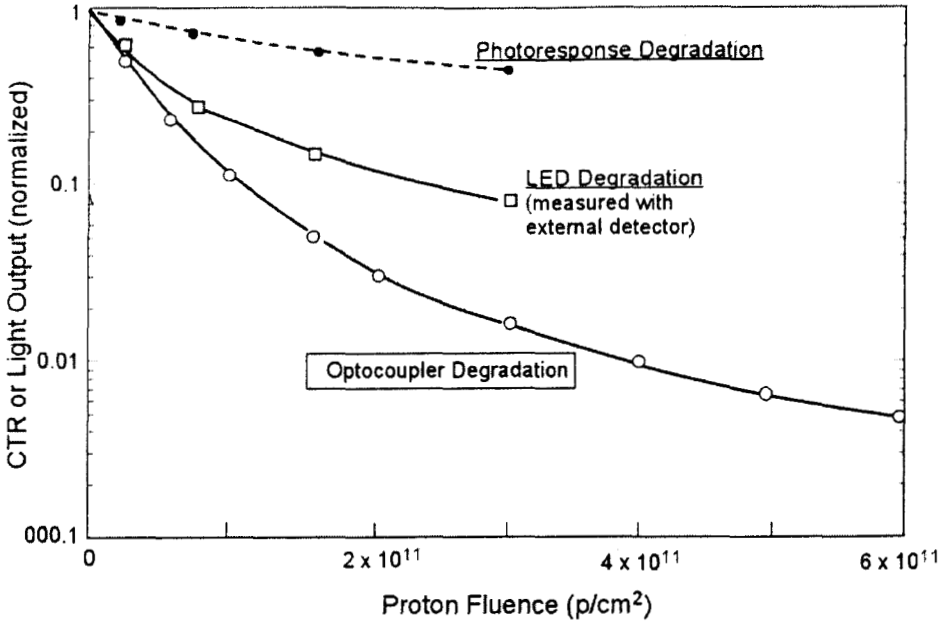



Figure 9-5. Degradation of the LED, photoresponse and gain of the components within the 4N49 optocoupler.

It is also useful to consider another basic type of optocoupler, the 6N140, that has also been widely used in space applications. The 6N140 uses an internal Darlington transistor. It requires a much higher input current at the LED in order to guarantee performance compared to the 4N49, but has faster response time. This device uses a 700 nm AlGaP LED that degrades far less than the amphoterically doped LED used in the 4N49. The shorter wavelength used in this optocoupler makes the phototransistor less sensitive to degradation, and helps to contribute to the improved radiation performance, even though this is a commercial part that is not designed to be hardened to radiation.

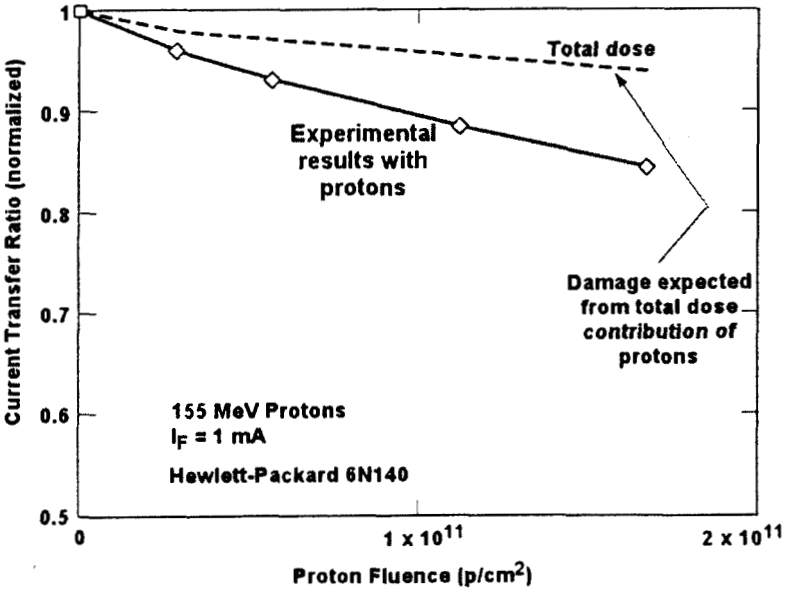



Figure 9-6 Degradation of the 6N140 optocoupler that uses a diffused LED with 700 nm wavelength.

Analog Optocouplers

Figure 9-7 shows how CTR for various forward current conditions is affected by proton irradiation [John2]. The LED in this particular optocoupler is an amphoterically doped LED, which is highly sensitive to proton damage. Prior to irradiation the maximum CTR occurs at about 0.5 mA. The CTR is considerably reduced for currents above that value, and normal practice is to operate above the peak in order to reduce the sensitivity of the CTR to temperature, gain and other variables, which is usually required of analog optocouplers. At low forward current, the CTR is strongly degraded by protons. This occurs because of two factors: the LED output drops, and the phototransistor gain depends on current. Thus, operating this type of optocoupler at low currents results in more degradation than expected from the LED.

If the optocoupler is operated well above the peak current, then the current dependence of the transistor gain reduces the sensitivity to LED drive, thus reducing the relative degradation. For example, if the optocoupler is operated at 2 mA the CTR decreases by about 30% at a fluence of 2 x $10^{10}$ p/cm$^2$. At 0.2 mA, the CTR decreases by about a factor of three at the same fluence.

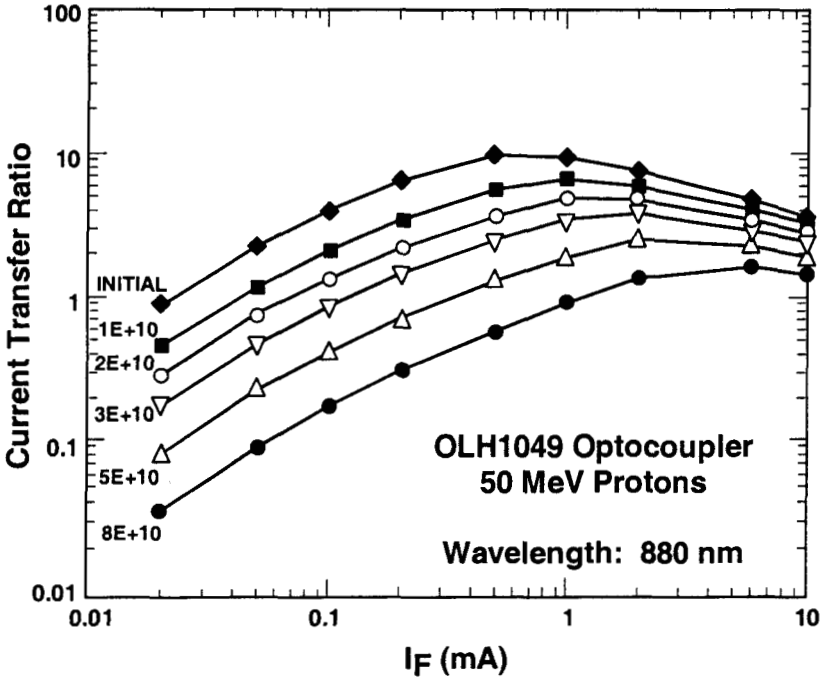



Figure 9-7. Proton degradation of an analog optocoupler with an amphoterically doped LED.

Just as for digital optocouplers considerable improvement in radiation performance can be achieved by using a different LED technology for analog optocouplers. Figure 9-8 shows proton degradation of an analog optocoupler with a double-heterojunction LED. It is degraded far less than the other optocoupler shown in Figure 9-7.

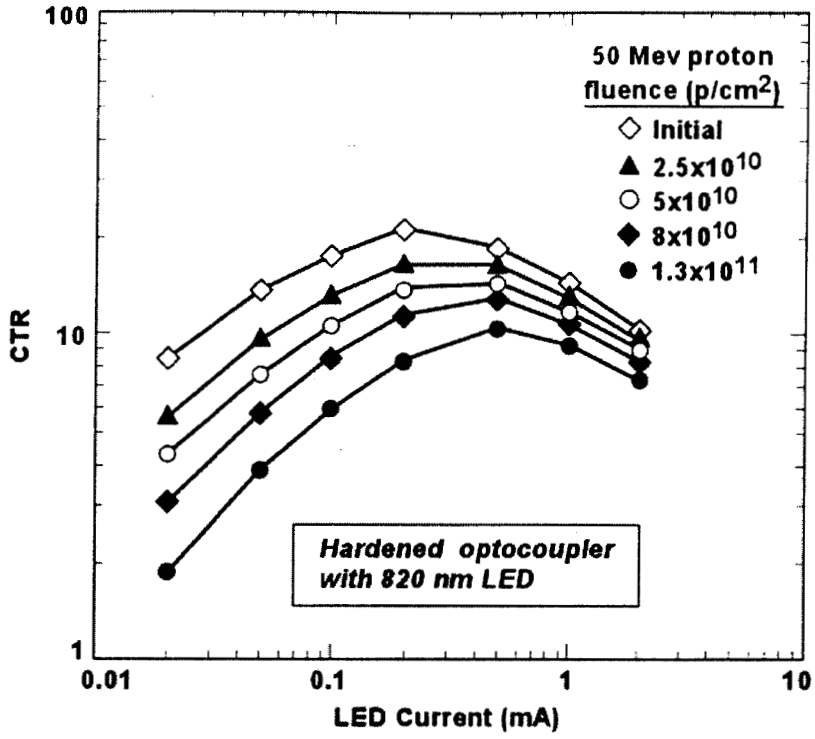



Figure 9-8. Proton degradation of an analog optocoupler that uses a double-heterojunction LED

### *C. Testing Issues*

Optocouplers generally use a combination of a III-V LED with silicon detectors and amplifiers. This creates an additional layer of confusion because the failure mechanism can be in either component, or for a combination of the two. For this reason it is particularly important to select an energy where the interpretation of NIEL is unambiguous, and 50 MeV protons are a good choice.

The construction of optocouplers must also be taken into account when selecting appropriate test energies. For devices with a sandwich construction care must be taken to ensure that the protons used for testing have sufficient range to penetrate the LED, ceramic substrate and other packaging material without degrading the energy to the point where the NIEL for either component is affected because of energy loss in layered materials.

Careful attention needs to be given to the way that electrical measurements are done between radiation levels. Injection-enhanced annealing can have a large impact on test results, particularly if high forward currents are used for some of the electrical tests. Pulsed measurements may be required to minimize annealing, as well as to keep from overheating the LED, which is affected by temperature (see Figure 9-2). As discussed earlier, special measurements of the transfer characteristics should be added to the normal set of measurements in order to measure the threshold behavior.

Physical factors are present in optocouplers that can also affect the radiation response and the operating margin. For example, the silicone coupling material in lateral optocouplers often contains bubbles that affect light transmission. The output of edge-emitting LEDs depends partly on the way that the edges are cleaved. These factors can lead to larger unit-to-unit variability in radiation response compared to conventional silicon semiconductors. Test sample sizes should be large enough to determine device variability for these classes of components.

## *10 - Solar Cells*

### *A. Construction and Electrical Properties*

Solar cells are essentially photovoltaic detectors that operate with no external bias. Unlike conventional semiconductors, they are designed to produce relatively high currents, with high efficiency. They are required to absorb light over a wide range of wavelengths -- from approximately 0.2 to 2 $\mu$m. Figure 10-1 shows a physical diagram of a basic crystalline silicon solar cell. It consists of a shallow n-region, diffused into a lightly doped p-substrate. The p-substrate has long lifetime, allowing light at long wavelength to be absorbed deep within the substrate. The diffusion length $L = \sqrt{D\tau}$, where D is the diffusion constant ($\approx$ ) and $\tau$ is the minority carrier lifetime.

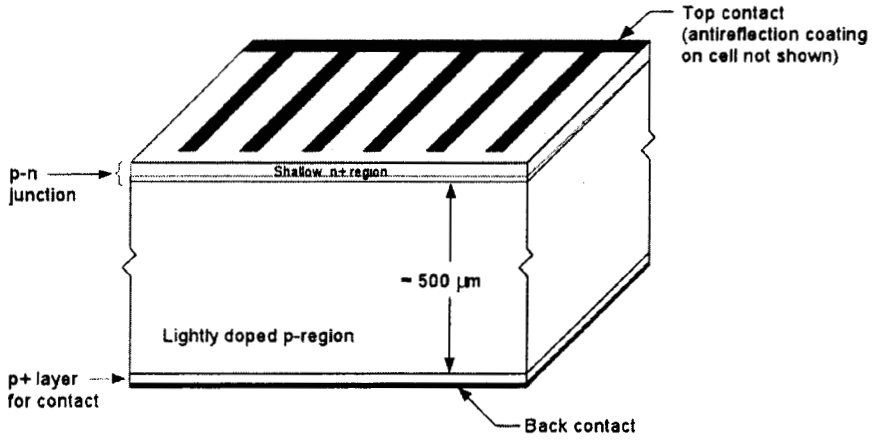



Figure 10-1. Diagram of a typical solar cell.

A low-resistance contact region is diffused on the top surface to provide a low-resistance path for current at the top of the cell and to allow ohmic contact to the metallization. The metallization is deposited as a lattice to minimize occlusion of the incident light. An antireflection coating is used at the top surface, designed to reduce reflection at the peak of the solar spectrum (about 0.6 $\mu$m). Although not shown in the figure, a thin cover glass is used on solar cells for space application to protect the cells. The I-V characteristics of a typical solar cell are shown in the inset in Figure 9-2, along with the solar spectrum. The cell develops a voltage above

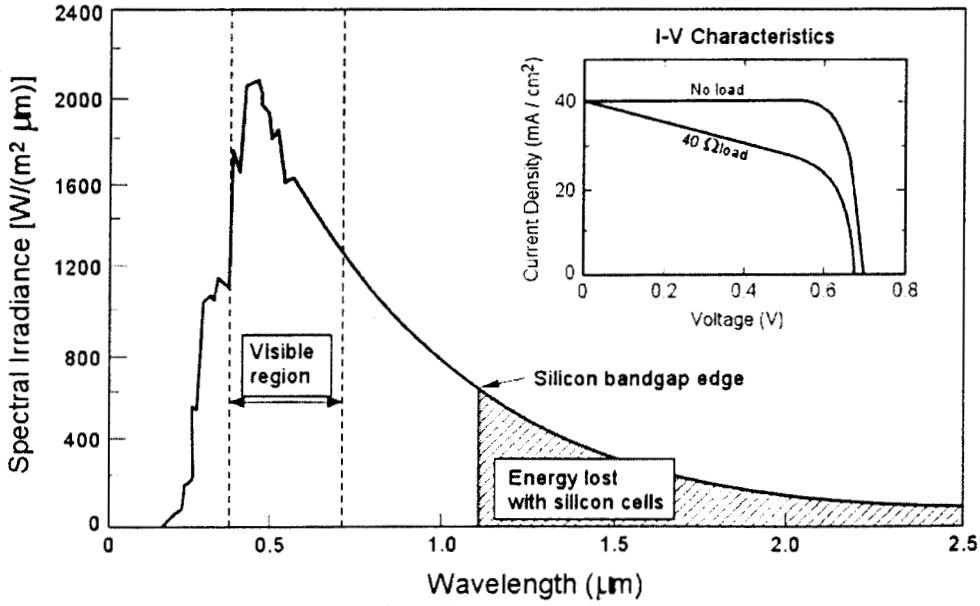



Figure 10-2. I-V characteristics of a typical solar cell

0.7 V with no load, but the voltage drops as the load current increases. The cell voltage and short-circuit current are basic parameters used to characterize the cell. Designs of basic silicon solar cells involve many tradeoffs. Low internal resistance is needed to reduce internal losses at high current, but that conflicts with the requirement to have high minority carrier lifetime and extended charge collection from deep regions in the cell. The solar spectrum extends from about 0.2 to 2 $\mu$m, a very wide range. Silicon does not absorb light beyond the absorption edge, approximately 1 $\mu$m, and consequently about 1/2 the energy of the solar spectrum is beyond the operating range of silicon solar cells. Considering the overall responsivity of silicon, the maximum efficiency of a basic silicon solar cell is about 29%.

There are many design features that can be used to increase solar cell efficiency, including (1) designing special lens assemblies, (2) using tandem cell designs that consist of a sandwich of two different types of cells, with the lower cell designed to efficiently absorb light at longer wavelengths that are beyond the absorption edge of the material in the top cell, and (3) using special concentrator assemblies that focus larger amounts of power on the cell. The cell efficiency is higher with increased power levels. Further details are provided in Reference Sol1.

### *B. Radiation Degradation*

Degradation of silicon solar cells is usually dominated by the decrease in minority carrier lifetime from displacement damage. The diffusion length is reduced, lowering the amount of energy that is collected at longer wavelengths. Figure 10-3 shows how the short circuit current of a typical n-on-p silicon solar cell is degraded by 10-MeV protons (the *de facto* standard energy for damage comparison in solar cell work because of the minimal amount of shielding). Degradation of this type of cell has been studied extensively, and the results agree closely with experimental results. Damage in solar cells is affected by annealing, and tests of solar cells must be done using special solar simulators that provide the operating power and total thermal heating of the cell that is expected in the actual environment. Because so little shielding is present, the mean energy of the proton distribution in typical space environments is significantly lower than 10 MeV. Consequently, the low energy protons from a single intense flare can degrade solar cell performance more than 30%.

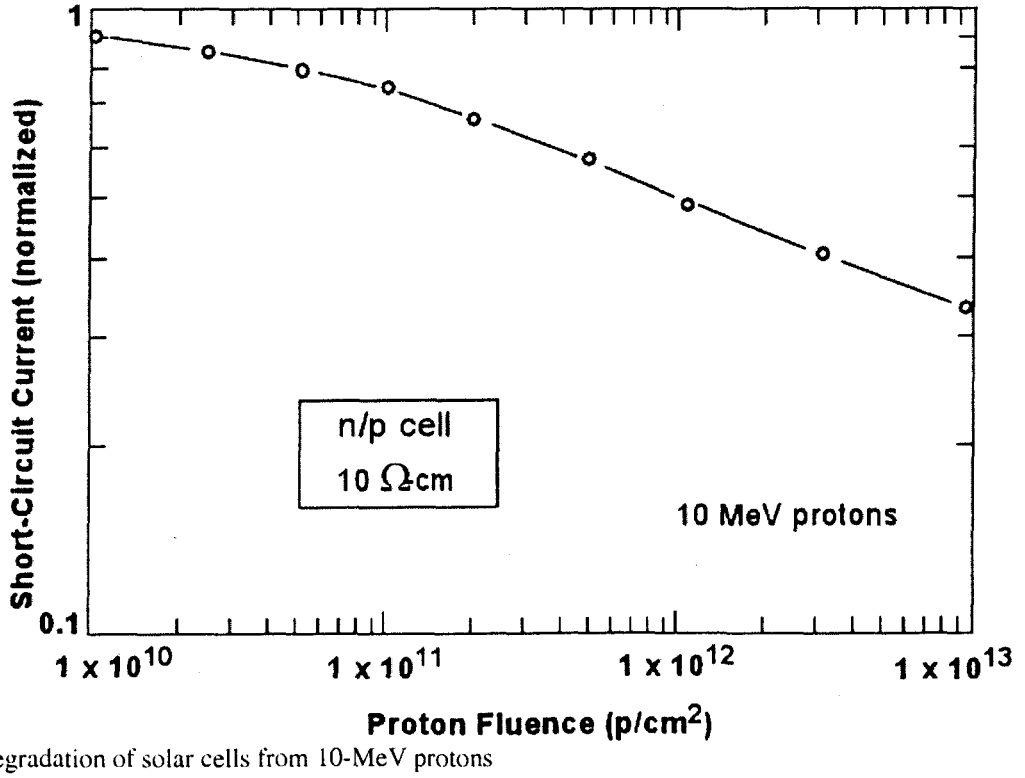



Figure 10-3. Degradation of solar cells from 10-MeV protons

Extensive work has been done to model solar cell damage, and closed-form equations to describe the damage have been developed for both silicon and GaAs cells [Sol2, Sol3]. Computer programs are available that can calculate the net degradation for a spectrum of electron and/or proton energies [Sol4].

Even though solar cell degradation has been studied extensively, there are limitations in the existing models, which basically consider only degradation from the reduced minority carrier lifetime . Those limitations became painfully obvious when solar cell panels failed abruptly on the NASA ETS-VI satellite [Sol5]. The reason for the failure was that the radiation level -- dominated by electrons in the earth's radiation belts -- was high enough so that carrier removal was a significant factor in cell degradation. Carrier removal effects were severe enough so that the p-region was converted to n-type material (recall that at high radiation levels the high levels of impurities "pin" silicon to n-type material because the impurity levels are not located at the center of the bandgap), with the result that the solar cells were no longer operational. Just before this catastrophic cell failure occurred the cell efficiency actually increased somewhat. Figure 10-4 shows how this degradation was modeled by Yamaguchi, et al. [Yamu1] for a lightly doped n-on-p cell that is 50 $\mu$m thick. The smooth line shows the predicted cell behavior when only lifetime degradation is considered. The other two curves show how carrier removal and changes in internal cell resistance affect the results.

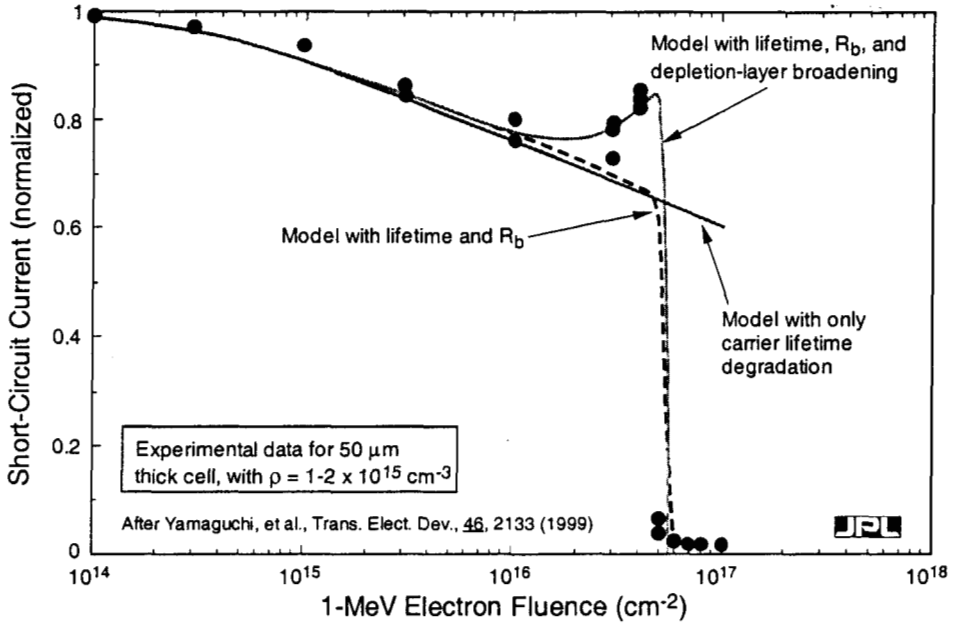



Figure 10-4. Anomalous degradation of lightly doped solar cells at high radiation levels

## *11 - Charge-Coupled Devices*

### *A. Conventional CCDs*

Charge-coupled devices are frequently used as imagers in spacecraft. The operation of a CCD is shown in the simplified diagram of Figure 11-1. Charge "packets" are induced by absorbed photons within each pixel. Information within the high-density array is transferred to the output by operating the device as a simple shift register, transferring the contents of each pixel laterally with a three-phase clock. This readout technique results in a very simple, open structure with high pixel density. However, its operation depends on the ability of the charge within each pixel to be shifted multiple times with minimum signal loss. This requires a high minority carrier lifetime within the array. One of the key CCD parameters is the charge transfer ratio (CTR), which is typically on the order of 0.9999 to 0.99999 for high-quality unirradiated cells. The charge transfer inefficiency, CTI, defined as (1 - CTR), is often used instead of CTR.

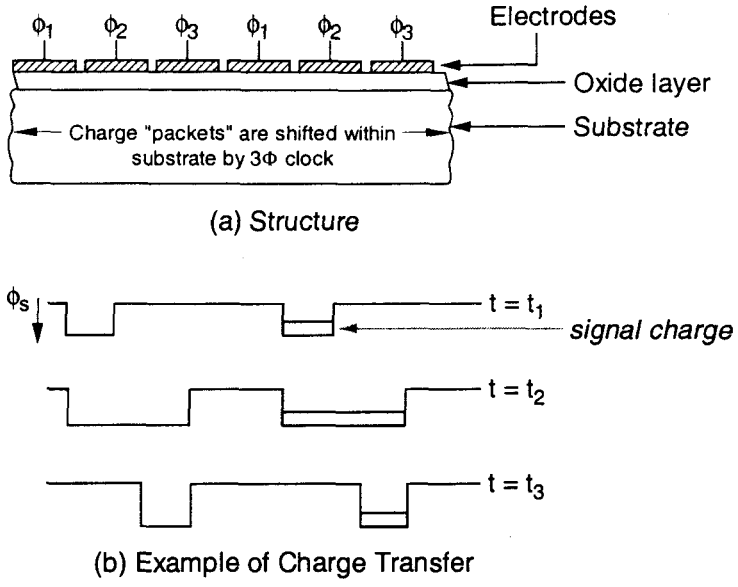



Figure 11-1. Simplified diagram of a basic CCD that uses a three-phase clock to transfer information.

CCDs are strongly affected by minority carrier lifetime degradation. Recent work has investigated an alternative p-channel CCD technology in contrast to the usual n-channel CCD. By changing the material type, the minority carriers are electrons, not holes, which are less affected by radiation damage. However, it is more difficult to fabricated p-channel CCDs, and consequently the initial CTR is not as high as for n-channel devices.

A comparison of radiation damage in n- and p-channel CCDs is shown in Figure 11-1, after Hopkinson, et al. [Hopk1]. The upper curve shows CTI for various internal signals for an unirradiated n-channel CCD. The lower set of curves show the performance of experimental p-channel devices after they are irradiated to 10 krad(Si); at that level the performance of typical n-channel devices is severely degraded. This illustrates the degree of improvement that is possible with different CCD technologies.

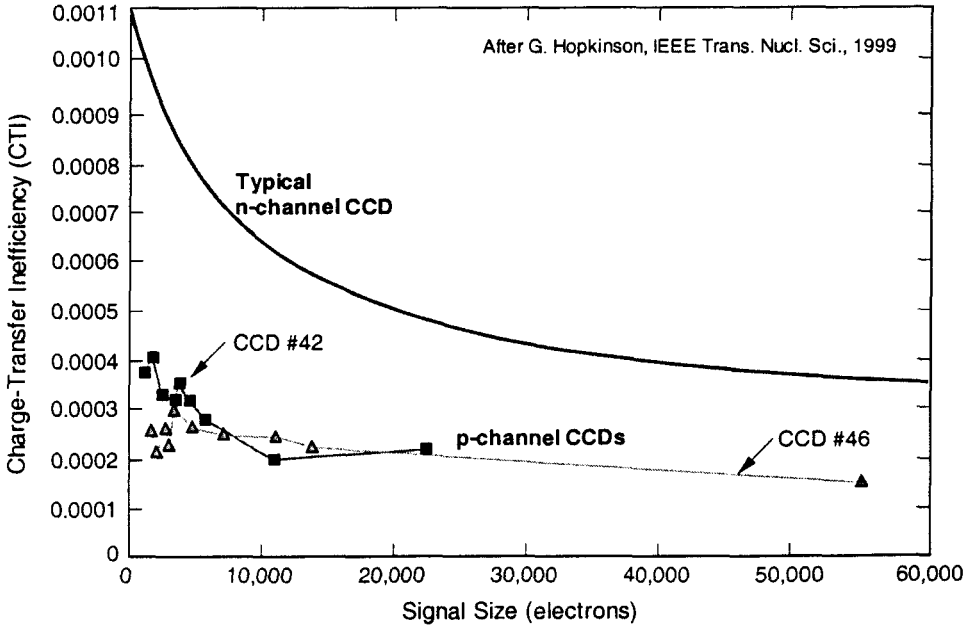



Figure 11-2. Performance of irradiated p-channel CCDs.

CCD pixels are small enough so that microscopic damage from the interaction of one or more protons can cause a small number of pixels to be severely degraded. Characterization of CCDs must take this into account, along with degradation of noise and sensitivity. Micro-dose damage effects in CCDs were reviewed by P. Marshall in the 1999 Short Course [Mars3], as well as by Hopkinson, et al. [Hopk2]. Numerous papers on CCD damage have been published within the last five years, and readers are referred to those references for more detail on CCD radiation effects.

### *B. Active Pixel Sensors*

An alternative approach for fabricating CCDs is integration of the electronics required for readout within the array. Such devices are called active pixel sensors (APS). The APS technology allows individual pixels to be addressed, eliminating the sequential series of charge-transfer events that are necessary for operation of conventional CCDs [Aps1, Aps2]. The resulting structure is far more efficient, and no longer depends on maintaining a very high minority carrier lifetime within the array. Thus, active pixel sensors are expected to perform far better in a radiation environment compared to conventional CCDs. They are also likely to be less affected by microscopic damage effects because of the "local" way in which information is accessed. Figure 11-3 shows a diagram of an active pixel sensor.

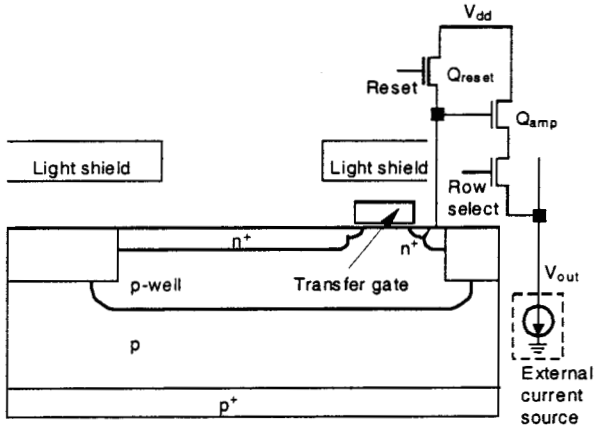



Figure 11-3. Diagram of an active pixel sensor.

Active pixel sensors are compatible with high-density CMOS processes, which are now available with very small feature sizes. This allows the array to be small enough to provide high resolution. A key requirement is advanced isolation (such as shallow trench isolation) that allows active pixels to be spaced close together. Active pixel sensors are far more complex than conventional CCDs, and this is an active area of development. Two papers in this year's conference will discuss radiation effects in active pixel sensors [Aps3, Aps4], and readers should consult those papers for up-to-date information about radiation performance of this technology. Active pixel sensors have also been developed that rely on direct heat absorption (bolometry) [Bol1]. They can be used at longer wavelengths, and are a promising new detector technology.

## *12 - Examples of Complex Failure Modes*

### *A. Optocoupler Failures in Space*

The first example of a complex failure mode occurred on the Topex-Posdeidon spacecraft, operating in a 1300 km five-year high-inclination orbit that is well within the inner portion of the van Allen proton belt. A basic optocoupler was used in several different applications, including direct control of engine thrusters used to maneuver the spacecraft. No radiation tests had been done on the optocoupler, and it was assumed that there were no significant radiation failure issues because the expected total dose level during the mission was less than 10 krad(Si). Although detailed information about the distribution of the total dose between electrons and

protons in the radiation belts was available, it was not provided as part of the abbreviated description of mission requirements and could only be obtained with considerable effort.

Several failures occurred in optocouplers after 2 1/2 years of operation [Swift]. The estimated proton fluence was approximately 2 x $10^{10}$ p/cm$^2$. The first failures that occurred were in circuits that transmitted the status of thrusters -- used to maneuver the spacecraft -- back to ground controllers during thruster operation. Fortunately the optocouplers used in the circuit that actually controlled thruster operation continued to operate. This was later found to be due to a much more conservative design practice for the optocouplers used in that particular circuit, which allowed the current transfer ration to degrade by a factor of ten before thruster operation would be affected.

The issues in this example are first that the designers of the spacecraft were unaware of the extreme sensitivity of the optocouplers that they used to displacement damage from protons. The second issue is that the environment was only specified as an overall total dose requirement, with inadequate awareness of the actual proton fluences that the spacecraft would encounter. The final issue is that the impact of the severe damage to the optocouplers was negligible in one application simply because the designers were ultra-conservative in their design implementation. This was a lucky happenstance. Because the potential of failure from the optocouplers was not recognized, it was very difficult to deal with the failures that actually occurred during the critical real-time operations that were involved in maneuvering the spacecraft.

### *B. Power Converter Failures During Ground Tests*

The second example is a hybrid power converter that was procured from a high-reliability manufacturer that had provided power modules for many previous spacecraft. This manufacturer, like many other hybrid manufacturers, considered all aspects of the design proprietary, and provided very limited information about either the design of the converter or the specific parts that were used to fabricate it. The intial parts list failed to include a linear optocoupler that was used to provide feedback from the output of the converter to the control electronics.

This issue was made even more complicated because the optocoupler (in itself a simple hybrid part) was procured from yet another manufacturer by the company that designed and built the hybrid power converter. Thus, we have "nested" hybrid manufacturers in the tree of parts used in the final product with limited awareness or control of either.

The optocoupler manufacturer procured LEDs from an outside vendor, and at one point changed the LED from a shorter wavelength double-heterojunction device (relatively tolerant to displacement damage) to an amphoterically doped LED that had better overall performance. The result was a decrease in the radiation hardness of more than one order of magnitude.

Test results for several hybrid converters made with this design are shown in Figure 12-1. The failure mode is lack of control of the output, with the output voltage gradually increasing towards the "raw" supply voltage. Note the very low levels of proton fluence at which these devices begin to degrade.

It is fortunate that this problem was identified before these devices were deployed. Normally proton tests would not be required for this type of part, particularly because the optocouplers were not included in the initial parts list provided by the power converter manufacturer.

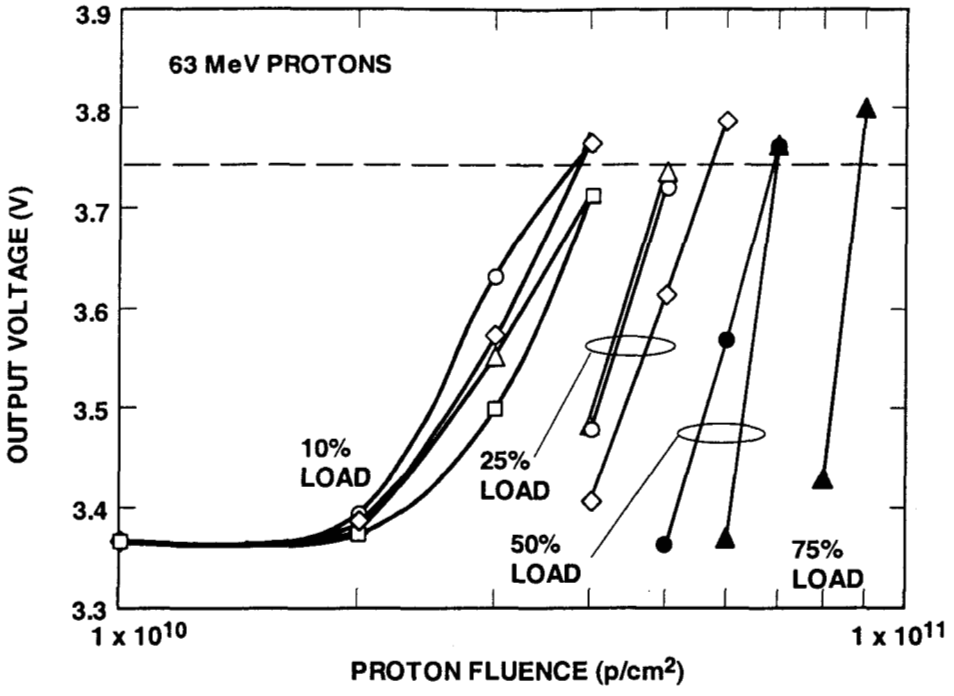



Figure 12-1. Failures of hybrid power converters during laboratory tests due to CTR degradation of linear optocouplers contained within the device.

## *13 - General Radiation Testing Issues for Optoelectronic Devices*

### *A. Radiation Sources*

Several different sources are available that provide high-energy protons. The highest energies are produced at cyclotron facilities. Van de Graaff accelerators are also available that can produce protons with energies up to 20 MeV, and may be less costly to operate. Table 13-1 lists three commonly used cyclotron facilities. All three provide steady-state proton beams with a circular beam area that is typically 5-10 cm in diameter. Operating costs for these facilities are nominally $500-$600 per hour.

Table 13-1. Comparison of Proton Accelerator Facilities

| Facility | Energy Range (MeV) |
|---|---|
| Univ. of California, Davis | 15 - 65 |
| Univ. of California, Berkeley | 50 |
| Univ. of Indiana | 65 - 200 |

High energy protons activate certain materials that are used in test hardware (particularly gold, copper and solder). The induced radioactivity can present a hazard to workers during experiments, particularly if tests are done at high fluences (> $10^{13}$ p/cm$^2$). Activation also affects shipping; it may not be possible to transport irradiated material to other facilities without waiting for a week or more for the induced radioactivity to die down.

A number of linear accelerators and Van de Graaff machines are available that can provide high-energy electrons, but they will not be discussed in detail because electrons are generally of less interest for

optoelectronic devices than high-energy protons (except for solar cells or other cases where very little shielding is present). When electron testing is needed energies of 3-5 MeV are typically used.

Cobalt-60 facilities (passive sources) are usually used to evaluate ionization damage. They are widely available at many aerospace companies, universities and government laboratories, but only simulate ionization damage, not displacement damage effects. We have not spent much time discussing ionization damage because displacement damage usually is the dominant effect for optoelectronic devices. Protons produce ionization as well as displacement effects, and there may be cases where ionization damage is really the dominant mechanism. However, tests with protons (or electrons where the environment is mainly from electrons) produce both ionization and displacement effects, so separate tests with gamma ray facilities are not required.

*B. Energy Selection*

Selection of energies for testing is a somewhat complicated problem. Usually the purpose of the test is to evaluate damage (essentially displacement damage) at a single energy to minimize cost and testing complexity, and then determine how the damage that is measured at that energy can be related to the damage produced in an actual space environment, where there is a wide distribution of proton energies. Usually this is done using the non-ionizing energy loss concept (NIEL), which was discussed at some length in Section III. One key point that was made is that NIEL calculations for GaAs do not always agree with experimental results for energies above approximately 80 MeV, and consequently energies used for proton testing should be below that value in order to avoid possible errors in interpretation of damage.

For protons, energies of 50 MeV are recommended because this is close to the mean proton energy of many earth-orbiting spacecraft, and the NIEL calculations agree reasonably well with experimental results for both silicon and GaAs. However, a lower energy is recommended for solar cells. This is appropriate because (1) there is much less shielding (at least on the front surface), and (2) nearly all of the archival results are done at 10 MeV. There are always special situations that may require different energies. Note particularly that some interplanetary missions may have very different energy ranges than earth-orbiting spacecraft.

Another factor in proton testing is the range of the protons. Some optoelectronic devices have considerable amounts of other material surrounding the critical devices, and the energy of the protons must be sufficient to allow them to penetrate to the active regions that are important for device operation without seriously degrading the energy. Note particularly the optocouplers with "sandwich" configurations where a thick III-V device and its associated substrate are placed over the silicon photodiode (see Figure 9-1).

Table 13-2 below shows the range for several proton energies assuming silicon (with density of 2.33; the range in other materials scales inversely with density). Note that GaAs has a density (5 .32)that is more than twice as large as that of silicon, so that the range of protons is considerably less in GaAs.

Table 13-2. Range of Protons with Selected Energy in Silicon

| Energy MeV | Range in Silicon ($\mu$m) |
|---|---|
| 200 | >20,000 |
| 100 | >20,000 |
| 65 | 18,000 |
| 50 | 8,610 |
| 30 | 5,220 |
| 20 | 2,580 |
| 15 | 1,585 |

Although we have not spent much time discussing testing with electrons, there are cases where electrons are a key factor in the environment. The solar cell results discussed in Section 11 are an example. Some interplanetary missions involve electrons with energies above 100 MeV. For earth-orbiting environments electron energies of 3-5 MeV are typically used. It is also possible to use 2-MeV electrons (conveniently available with some sources). The energy of electrons is generally less important than energies selected for protons because NIEL for electrons depends much more gradually on energy, provided the energy is above the displacement damage threshold (about 150 keV). Table 13-3 shows the range of electrons of various energies in silicon.

Table 13-3. Range of Electrons at Selected Energies in Silicon

| Energy MeV | Range in Silicon ($\mu$m) |
|---|---|
| 1 | 2,800 |
| 2 | 4,700 |
| 3 | 6,900 |
| 10 | 17,200 |
| 100 | 138,000 |

Most laboratory tests of optoelectronic devices is done with the device normally incident to the beam direction. This makes it relatively straightforward to make sure that the range of the particles that are being used is adequate to go through packaging material, lenses, coatings and any other material that is present in the structure. It may be necessary to do complex transport calculations to determine the effective energy in the real environment for cases where a great deal of extra material is present. For example, solar cells have very little shielding on the top surface, but the energy of electrons and protons that go through the device at angle (or from the back of the cell) will be different. Similar issues are important for other applications, including CCDs or detectors that are behind baffles and shields. This does not necessarily affect the way that radiation tests are done, but it does affect the interpretation of the test results.

### *C. Single-Event Upset Testing*

Although single-event effects are not addressed in this part of the course, a brief discussion on single-event testing has been included for completeness that examines some of the special issues that have to be dealt with when optoelectronic devices are subjected to single-event testing. Before beginning this discussion, recall that SEE effects are usually described in terms of linear energy transfer (LET), with units of MeV-cm$^2$/mg. Galactic cosmic rays and solar flares have a continuous distribution of LET values, up to an LET of about 100 MeV-cm$^2$/mg [ref], but the number of particles at high LET falls rapidly as the LET increases. There is an abrupt drop in the LET distribution at about 30 MeV-cm$^2$/mg (the "iron" threshold) that is of great practical interest. Because relatively few particles are present with LET's above 30 MeV-cm$^2$/mg, devices with threshold LET values above that value† are relatively immune to single-event upset. On the other hand, some devices (such as DRAMs) have threshold LET values of 1-2 MeV-cm$^2$/mg, and are extremely sensitive to single-event upset effects.

Two types of single-event tests will be considered, as depicted in the simple diagram in Figure 13-1:

(1) Tests with heavy-ions that have very limited range (typically 30-50 $\mu$m) and must be done in a vacuum chamber; and

(2) Tests with high-energy protons that have much longer range, but can affect devices either *directly* from proton ionization, or *indirectly*, via a nuclear or kinematic collision with a lattice atom. Direct ionization produces ionization tracks with low charge density (LET approximately 0.1 MeV-cm$^2$/mg), but some optoelectronic devices are affected by direct ionization. The indirect process produces a recoil atom with higher effective LET, but a relatively short track length (a few $\mu$m).

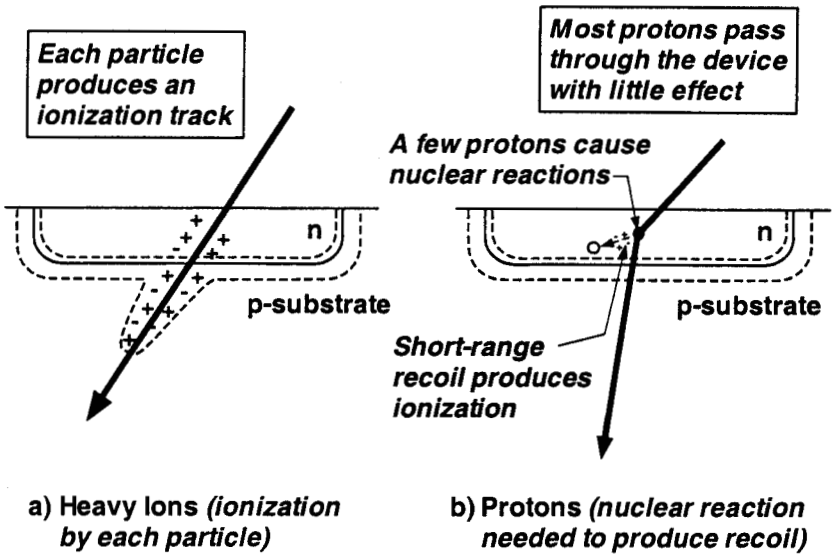



Figure 13-1. Diagram showing direct interactions from high-energy cosmic rays and indirect interactions from protons through an intermediate nuclear reaction.

Physical Factors

The first problem is that of recognizing that unlike silicon integrated circuits, the active region of many optoelectronic devices is located far below the surface of the device. For example, the active region of a typical LED is 50-100 $\mu$m below the top region. This presents a major problem for heavy ion tests because it is difficult to find a source with sufficient penetration depth. Most facilities are limited to ranges below 100 $\mu$m for ions with LET values above approximately 20 MeV-cm$^2$/mg [Koga1]. Fortunately, single-event effects in LEDs and laser diodes are generally of minor importance, and usually one is more concerned about SEU effects in amplifiers and detectors.

Optocouplers are an example where physical construction is very important. For optocouplers with the LED array placed over the photodiode it is generally not possible to do SEE testing with heavy ions because of occlusion by the LED. One way to deal with this is to disassemble the device, removing the LED array, allowing unobstructed access to the photodiode and amplifier chip. Even after the LED assembly is removed, a layer of optical coupling compound is still present on the top of the silicon chip that has to be removed with a solvent. The silicon chip that remains cannot be tested as an optocoupler, but is essentially equivalent to the optocoupler in the "off" state.

Electrical Requirements

Many optoelectronic components (including optocouplers) are relatively slow devices that are designed to work with load resistances of several kilohms or more. They cannot drive terminated 50-ohm cables, and it is generally necessary to use an active line driver in order to monitor output signals. The output response of optocouplers is strongly affected by resistive and capacitive loading, and the test conditions must closely mimic the actual application.

Example: Heavy Ion Results for a Digital Optocoupler

An example of test results for heavy-ion test of a digital optocoupler is shown in Figure 13-2 (the LED assembly was removed before the tests were done). There are several subtleties that are not obvious at first glance. With heavy ions, this device is extremely sensitive to upset effects. The threshold LET of approximately 0.3 MeV-cm$^2$/mg is well below the level that even highly sensitive dynamic memories exhibit upset. However, unlike a digital circuit there is a wide variation in the pulse amplitude and pulse width. The nature of the output depends where the ion happens to strike the device. Thus, the cross section has to be defined in terms of specific criteria for output amplitude and pulse width, as well as loading conditions.

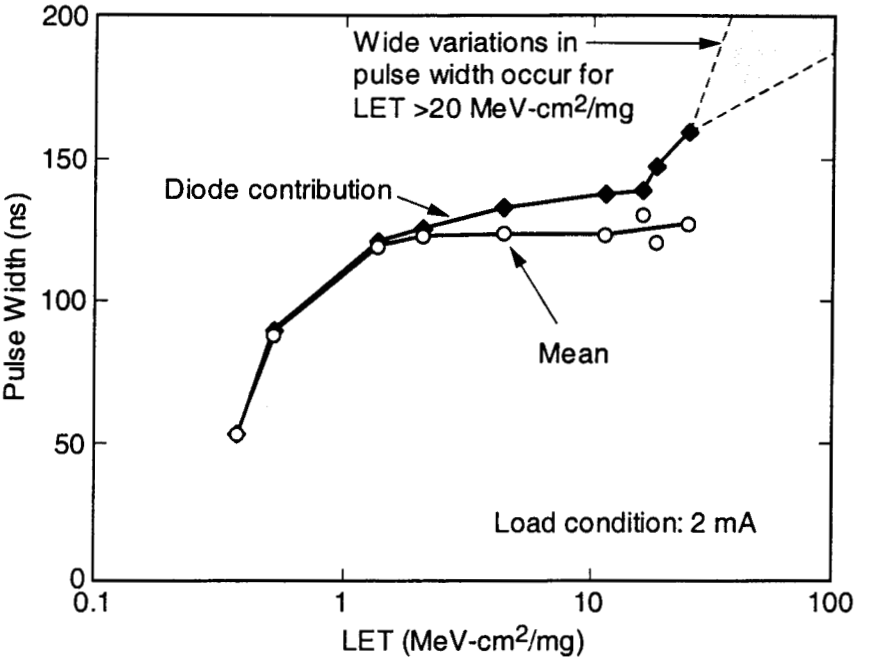



Figure 13-2. Variation of mean pulse width from a digital optocoupler at various LET values.

Near the threshold the pulse width is extremely narrow, gradually increasing as the LET of the ions used for testing is increased. The pulse width gradually increases to about 130 ns, and most of the pulses remain at about that pulse width until the LET exceeds 10 MeV-cm$^2$/mg. However, prior to the point where the pulse width begins to increase a small number of pulses are apparent with considerably lower amplitude. This causes the mean pule width to decrease slightly, and that is due to the gradual contribution of a second mechanism, the response of the high-gain amplifier. Once the amplifier is fully turned on the cross section and the nature of the pulse widths that are observed change radically, as shown in the figure.

Direct Ionization Effects in Optocouplers

The photodiodes used in optocouplers have a large diameter compared to that of most components, and this makes it possible for direct ionization from protons to be a factor in their response. Because of the large diameter, far more charge is collected when the proton passes diagonally through the photodiode compared to the charge generated at normal incidence. This was first observed by LaBel, et al. [Label]; the cross section increased by nearly an order of magnitude when they carried out tests of optocouplers at high angles.

A more thorough investigation of the effect of direct ionization was done afterwards [John3] that used a much wider range of proton energies. In this case it was not only necessary to remove the LED assembly, but also to grind down part of the side of the package in order to allow protons with lower energies to reach the surface of the die, which was recessed somewhat in the package. The results of these tests are shown in Figure 13-3. Note that at lower proton energies direct ionization has increased the cross section by about three orders of magnitude. This is a sufficiently large increase to affect the upset rate in the real environment.

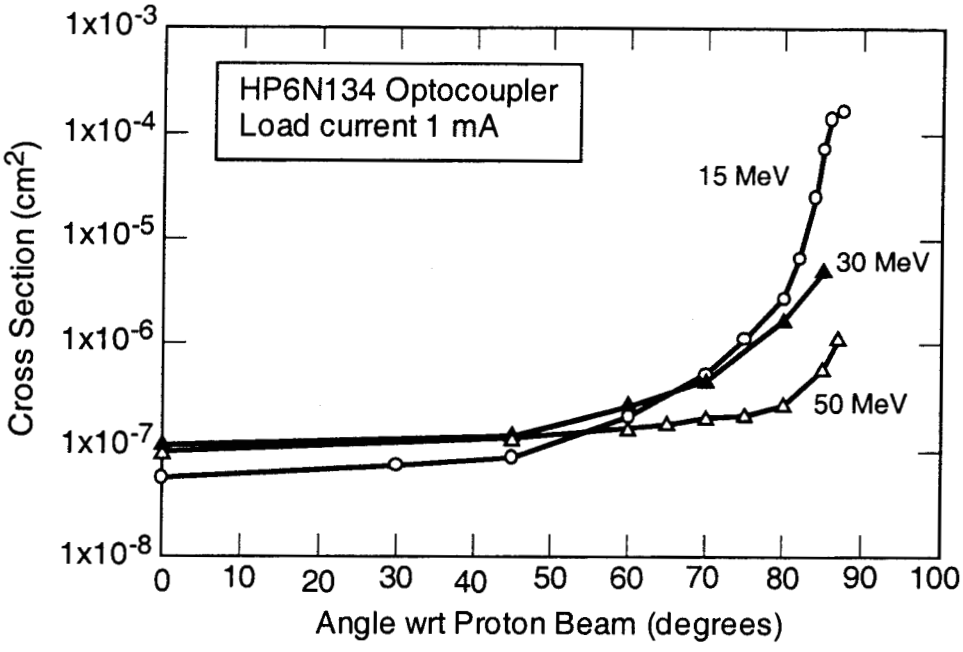



Figure 13-3. Dependence of upset cross section on incident angle for an optocoupler for various proton energies.

Note that even though low energy protons in an experiment are partially shielded by the device package, protons with the continuous distribution of energies in the actual environment are simply shifted down in energy. Thus, although 15 MeV protons will not penetrate the package during a test at a single energy, substantial numbers of protons in that energy range will reach the active part of the device in the real environment because protons with somewhat higher energy will lose part of their energy when they pass through the package. Thus, although one might initially think that interference in this type of experiment from self-shielding by the package makes the effect inconsequential, that is not the case when one considers the net effect of shielding on the distribution of proton energies within the spacecraft (and within the device).

The data in Figure 13-3 required an extensive effort that is well beyond that normally expected for routine characterization. of this type of part. An alternative way to estimate whether indirect ionization is important was develop in Reference [John3] using relatively straightforward tests with laboratory alpha particle sources. That method can be used as an initial screen to determine whether the cost and difficulty of detailed tests at angle is needed for specific devices.

## *14 - Summary*

This part of the 2000 NSREC Short Course has discussed many aspects of optoelectronic devices, which are not as well understood as more conventional semiconductor devices. The radiation response of optoelectronic devices is usually dominated by displacement damage effects that can become important at relatively low radiation levels. This is often overlooked because most other types of electronic components are relatively unaffected by displacement effects until much higher radiation levels are reached.

A considerable amount of material was presented on the design of optoelectronic devices. This was done for several reasons, in particular because some types of new optoelectronic structures are very different from

their older counterparts. It is usually necessary to have a basic understanding of these devices in order to evaluate radiation test data, or to plan radiation tests. LEDs and laser diodes are examples where device technology has evolved in several different directions, making it particularly difficult to select devices for use in space.

Optocouplers are good examples of the way that different effects interact to produce a complex interdependence of failure modes, requirements, and use conditions. Some optocouplers are extremely sensitive to proton displacement damage and have actually failed operationally in space.

There are many evolving optoelectronic devices -- vertical cavity semiconductor lasers, active pixel sensors, and others we have not discussed -- that are likely to be seriously considered for future space applications, particularly because of the need to decrease the size, weight and cost of spacecraft. New ways of using optoelectronics for optical interconnects or as special dedicated integrated optic devices are being developed, and it is likely that radiation effects in these structures will be an interesting topic during the next decade. There are a number of more exotic structures -- including the use of porous silicon light emitters [ref] or avalanche emission from silicon [ref] -- that have been considered to allow direct integration of light sources with silicon-based electronics. The intent of this part of the course is to provide the necessary background to understand and appreciate the mechanisms and principles that affect optoelectronics.